AZ EMBER ÁLTAL FELVETT INFORMÁCIÓ MÉRÉSE A SZUBJEKTIV VALÓSZINÜSÉGELOSZLÁS MEGVÁLTOZÁSA ALAPJÁN

Doktori értekezés

Szekeres István

Nagykőrös,

1981.

T A R T A L O M

B E V E Z E T É S

Az információ igen általános, napjainkban sürün használt fogalom, mely fontos szerepet tölt be az egyes tudományokban, igy a matematikában, a pszichológiában és a pedagógiában is. Ez a kifejezés azonban nem ugyanazt a fogalmat takarja a felsorolt tudományokon belül, de kapcsolatba hozhatók egymással. Törekvések folynak az irányba, hogy az időrendben először egzaktan megfogalmazott matematikai definiciót a pszichológiai és a pedagógiai értelmezésben is felhasználják. Természetesen a pszichológiai és a pedagógiai értelmezés már nem térhet el egymástól, hisz az információfelvevő mindkét esetben ember, egy szubjektum, igy az alább leirt módszerek mindkét tudományon belül alkalmazhatók, ha rövidség kedvéért csak az egyiket emlitjük is.

Az információ matematikai módszerekkel történő vizsgálatát bizonyos /először a hiradástechnikában fellépő/ gazdaságossági problémák inditották el. Alapprobléma az optimális kódolás kérdése volt, mely az információelmélet kialakulásához vezetett.

Matematikában az információ egy eseményrendszer adott eseményének bekövetkeztéről való tudósitást jelent, melyhez a

$$\sum_{i=1}^{n} Pi \log \frac{1}{Pi}$$

definició alapján számot / mint a kisérletnek önmagára vonatkozó információ értékét/ rendelnek, igy mód adódik az információ mérésére.
/A fenti definicióra, annak magyarázatára még visszatérünk./

Az ember által felvett információ mérése a közelmultban a pszichológia /vele együtt a pedagógia/ egy sarkalatos problémájává vált. A mérés szükségessége több irányból is felvetődött. Itt elegendő csupán az "ember-gép" rendszerrel, vagy némely tanuláslélektannal foglalkozó tanulmányra hivatkozni.
/ld.: [10], [19], [20], [6], [22] és [15] irodalmak/.

Emlitettük, hogy az "információ" kifejezés más és más tartalommal bir az egyes tudományokban, Már a pszichológián belül sem egyértelmü e fogalom, szükebb és tágabb értelemben használják, ÁDÁM György / [2] / például már az ingerületi folyamatot információfelvételnek tekinti, és igen érdekes módon hasonlitja össze az általános információs csatornával. Ilyen értelemben a kódolást a receptorok végzik, az idegrostok a jelzések továbbitására szolgáló berendezések, a dekódolás pedig az agyban történik.

Valóban, ha az információfelvételt részletesen akarjuk vizsgálni, az analizátorok müködésének elemzésével kell kezdenünk, hisz ezek biztositják kapcsolatunkat a külvilággal, ahonnan az információ származik. A módszer megválasztása a céltól függ, esetenként felesleges az ilyen részletekbemenő vizsgálat, hisz távol vagyunk még attól, hogy az ember információszerzési folyamatát ilyen részletességgel végig nyomon tudjuk követni. Sokszor meg kell elégednünk a folyamat felületesebb vizsgálatával, olyan értelemben, hogy csak a kiindulási inger és a végső reakció kapcsolatát vizsgáljuk, a közbenső belső biológiai és pszichológiai folyamatok ismerete nélkül. Ilyen értelemben beszélnek a kutatók információfelvételről, ha például egy tárgy észlelése utján felvett információról van szó. Nyilvánvaló, hogy az igy értelmezett információfelvétel sokkal komplexebb folyamat az előbbinél, egy tárgy észlelése igen sok idegpálya által közölt információ eredménye. Mi is ez utóbbi értelemben használjuk az információfelvétel fogalmát. Ezt a módszert ITELSZON [22] makromódszernek nevezi, és a rádióhoz hasonlitja, amelynek belsejéről sok ember nem tud semmit, de tudja, honnan van az adás, és érti, mit mond a rádió. Ez a hasonlat a mi esetünkben nem tükrözi elképzelésünket, ugyanis mi már csak az agyban feldolgozott információval foglalkozunk, ami eltérhet a közléstől, hisz az érzékszervek, az idegpálya zajos csatornaként kezelhető.

Az információfelvétellel kapcsolatban részletekbemenő kutatások folynak /különös figyelmet szentelnek a kutatók az emberrel közölhető maximális információmennyiség felmérésére és az ember "csatornakapacitására"/, ugyanakkor a felvett információ mérése még mindig problematikus, nem egyöntetü az egyes kutatók esetében, "még mindig nincs világos válasz arra az alapvető módszertani kérdésre, hogyan kell mérni az idegrendszer által felvett és feldolgozott információt" /LOMOV 1969. 167. old./. Általában az információelméletben kidolgozott definiciókat és képteleket alkalmazzák az ember által felvett információ mérésére. Itt is előfordul azonban az a /matematikai módszerek alkalmazásakor gyakori/ hiba, hogy az egyes, matematikában definiált fogalmakat, matematikai összefüggéseket mereven alkalmazzák, tulzottan leegyszerüsitik a lelki folyamatokat a tárgyalás egyszerüsége érdekében, igy egyrészt nem használódik ki eléggé a matematikai módszerekben rejlő lehetőség, hisz e módszerek bonyolultabb folyamatok leirására is alkalmasak, másrészt nem irja le annak a folyamatnak a lényegét, amire kidolgozták.

Az információfelvétellel foglalkozó irodalom többnyire egy tárgy, vagy jelenség észlelése által felvett információ mérésével foglalkozik, ami általában két módszer szerint történik. Az egyik az, hogy megállapitják a szóbanlevő tárgy előfordulási valószinüségét /p/, és ezután az információelméletből ismert $\log \frac{1}{p}$ formula alapján számitják a felvett információ értékét.

E módszerrel kapcsolatban felmerülő problémák /pl: mennyire egzaktan tudjuk felmérni a tárgy előfordulási valószinüségét/ ellenére helyes az a módszerben rejlő gondolat, hogy a ritkán előforduló tárgyak, jelenségek észlelése nagyobb meglepetést okoz, nagyobb információt nyujt a szubjektum számára. A másik módszer lényege, hogy a tárgyak egyes tulajdonságai által közölt információkat külön-külön mérik, végül ezeket összegzik. A tárgy egy tulajdonsága /pl: szine/ által közölt információt ugy definiálják, hogy felmérik az egy ember által differenciálható elemi tulajdonságok számát /pl: hány szint tud megkülönböztetni/ és ezeket az elemi tulajdonságokat teszik meg az abécé alapjául. Ha például egy ember 500 szint tud megkülönböztetni, akkor egy szin észlelése log 500 információt hordoz.

LOMOV mindkét módszert birálja, s másodikkal kapcsolatban rámutat arra, hogy az összinformáció értéke nem mindig egyezik meg a részinformációk értékeinek összegével. Cáfolja, hogy egy tárgy észlelésekor kapott összinformáció megegyezik annak hosszának, szélességének, alakjának, szinének, stb. észlelésekor kapott információk összegével. Módszerünk egyik előnye éppen az lesz, hogy ez a probléma természetes módon oldódik meg, látni fogunk példát arra, hogy két információ összege adja az összinformáció értékét, de ellenpéldát is bemutatunk.

Az ember által felvett információ mérésére használt módszerek /amikor csak a matematikai valószinüségekkel számolnak/ egyik hiányossága, hogy nem tudják megfelelően figyelembe venni az egyéni adottságokat, általában beszélnek arról, hogy egy konkrét tárgy észlelése mekkora információt tartalmaz. Márpedig nem mindegy, hogy azt a tárgyat ki észleli. Nyilvánvaló például, hogy az autó alkatrészeinek megtekintése a szakember számára több információt szolgáltat, mint egy laikus esetében.

A fenti problémák csak halványan érzékeltetik annak szükségességét, hogy az ember által felvett információ mérését más uton közelitsük meg. E dolgozat feladata, hogy az ember által felvett információ mérését pontosabbá tegye, a méréshez szükséges matematikai fogalmakat a pszichológia és a pedagógia számára használhatóbb formában fogalmazza meg, általában igyekszünk olyan módszert kidolgozni, ami a pszichológia sajátosságainak jobban megfelel. Ezzel nem akarjuk azt állitani, hogy a leirásra kerülő módszer olyan tökéletes, hogy lezárja a problémakört, bizonyosnak látjuk azonban, hogy az eddigi módszerek lényeges általánositását, továbbfejlesztését sikerül elérnünk, abban reménykedünk, hogy igy a valóságot jobban közelitő modellhez jutunk és az információelmélet adekváltabb módszerévé válik e kérdéskör vizsgálatának.

E dolgozatban nem kötődünk egyik emlitett módszerhez sem, az ember által felvett információ mérését más irányból közelitjük meg. Az általunk leirt módszer is módot ad arra, hogy egy tárgy észlelése utján felvett információt mérjük, de nem ebből indulunk ki.

Messzemenően figyelembe vesszük azt a VASZKÓ Mihály [19.] által is megfogalmazott tényt, hogy az információ bizonytalanságot feloldó közlést jelent. L.N. LANDA [23.] ezt ugy fogalmazta meg, hogy a bizonytalansági szituációt megszünteti, vagy csökkenti az információ. Ez történhet más ember által való közlés alapján, vagy magának az embernek saját megismerési tevékenységével. Ez utóbbi definició az előzőnek általánositása,hisz ez nem követeli meg a bizonytalanság teljes feloldását. A mi definiciónk ez utóbbit még tovább általánositja.

Számunkra mindenképpen fontos, hogy a definició alapján matematikai módszert alkalmazhassunk. Nem áll ez ITELSZON definiciójára, miszerint "az információ valami, amit a közlések tartalmaznak".

Egy ember által felvett információ értékének definiálását tüztük célul magunk elé. Ezzel egy lelki folyamat eredményét kivánjuk számmal jellemezni. Az ilyen irányu mérés nem kis elvi problémát jelent, sokszor vitára ad alkalmat, hisz vannak akik elvetik, vagy legalábbis erősen vitatják ennek létjogosultságát. Hely és felkészültség hiányában nem kivánunk e kérdéssel részletesen foglalkozni, csupán a gyakorlatra hivatkozva indokoljuk alábbi munkánk alkalmazhatóságát. Meglehet, hogy egyetlen szám nem sokat mond egy szubjektumról, sőt vannak olyan tulajdonságok, melyeket valóban nem lehet még mérni /bár ez a tudomány fejlettségétől is függhet, hisz a fény hullámhosszának létezése ismerete előtt ki gondolta volna, hogy a szineket sorba lehet szedni/, mégis mérünk emberi tulajdonságokat, amiket a gyakorlat kényszerit ránk, ezek szükségességét és létjogosultságát nem lehet vitatni. Itt gondolhatunk a tanulók osztályozására, vagy a pulzus és a vérnyomás mérésére, ami alapján egy ember idegállapotára is következtethetünk, de ide sorolhatjuk a különböző intelligencia teszteket is.

Bár elfogadjuk azt a MANNHEIM idézetet, melyet BÉKÉS Ferenc [4.] választott könyve mottójának /"Nem biztos, hogy a társadalomtudományokban azok a legfontosabb összefüggések, amelyeket mérni tudunk"./, mégis hisszük, hogy bizonyos pszichológiai fogalmak mérése segiti e tudomány fejlődését. Különösen indokoltnak és fontosnak tartjuk az információ értékének mérését, mert e problémával /mint már említettük/ a pszichológusok és a pedagógia müvelői igen sürün találkoznak.

A mérés problémája szervesen kapcsolódik a matematika alkalmazásának problémájához. A mérések eredményeivel való számolás ugyanis a matematikai törvények alapján történhet. Itt szeretnénk kihangsulyozni, hogy a matematika alkalmazása nem azonos az alapmüveletek és százalékszámitási feladatok elvégzésével. E kérdésben is megoszlanak a vélemények. ARATÓ Mátyás [3.] például azt irja, hogy a "matematika ma még nem képes arra, nincsenek meg hozzá az eszközei, hogy bonyolult gondolkodási, társadalmi folyamatot vagy gazdasági folyamatokat le tudjon irni". E véleménnyel olyan értelemben értünk egyet /valószinü ő is igy érti/, hogy még nem törekedhetünk teljességre. A matematika viszont már ért el nagy eredményeket, amelyeket /legalábbis egy részét/ viszont fel lehet és fel kell használni a társadalomtudományokban, ha csak részfeladatok megoldására is.

Természetesen helytelenitjük a matematikai módszerek erőltetett alkalmazását. Hangsulyozzuk viszont azt a véleményünket, hogy a matematika alkalmazásakor az ott megszokott módszert kell követni. Ha egy matematikai fogalmat más tudományra alkalmazzuk, akkor a matematikai elnevezést és jelölést célszerü megtartani. Nagyon helytelen az a gyakorlat, hogy valamely tanulmány közöl egy-két képletet, de az olvasó elől rejtve marad annak eredete. Igy a további kutatás számára /a szerzőt leszámitva/ elveszett, hiába adják meg a képletben szereplő betük jelentését, a képlethez vezető ut megmutatásának hiánya azt vonja maga után, hogy az olvasó nem ismeri meg az alkalmazhatóság feltételeit, számára az esetleges továbbfejlesztési lehetőség nincs megadva. Erre azért tértünk ki, hogy ezzel is indokoljuk azt a közlési módszert, amit alkalmazni fogunk e dolgozatban.

E dolgozat közlési módszerét illetően bizonyos kettősség fog érvényesülni: dolgozunk integrálokkal, ugyanakkor egyszerübb számolási müveleteket is magyarázunk. Ennek oka, hogy e dolgozat témája a pszichológia és a pedagógia egy igen izgalmas problémája, de ezt matematikai módszerekkel dolgozzuk fel. Igyekszünk tehát ugy leirni mondanivalónkat, hogy a matematikában kevésbé járatos olvasó is megértse lényegét, de a matematikai lényeget sem szeretnénk elbagatellizálni. Feltételezzük, hogy az olvasó, netán e témát továbbfejlesztő pszichológus kikéri egy matematikus véleményét, aki viszont akkor érti meg könnyebben a dolgozat tartalmát, ha az a matematika nyelvén van leirva. E kettős feltételnek ugy kivánunk eleget tenni, hogy ragaszkodunk ugyan a matematikai jelöléshez, elnevezésekhez, de ezeket részletesen magyarázzuk. A szereplő fogalmakat, megállapitásokat több konkrét példa bemutatásával tesszük még érthetőbbé. Ragaszkodva az előbb leirt véleményünkhöz, szerepelni fognak a szükséges matematikai levezetések, de ezek átugrása sem okoz problémát, mert a végképletet a következő példákból ugy is megérti az olvasó. A matematikai levezetéseket értő olvasó számára viszont néhány példa felesleges magyarázkodásnak tünhet. Sajnos előfordulnak olyan részek, amelyeknél a matematikai apparátus magyarázata e dolgozatban lehetetlen, de ezek elhagyása /bár többnyire lényeges mondanivalót takarnak/ nem akadályozza a további fejezetek megértését.

Megjegyezzük még, hogy a Magyar Pszichológiai Szemle 1976/5. számában megjelent [15] dolgozatunkban is találhatók olyan konkrétan elvégzett, részletesen leirt számitások, melyek átnézése elősegiti az itt alkalmazott számitások megértését.

I. A matematikai és a szubjektiv valószinüség-eloszlásról.

E dolgozat alapja a Magyar Pszichológiai Szemle 1979/3. számában megjelent „Szubjektiv valószinüségeloszlás különböző tipusu feladathelyzetekben" cimü dolgozatunk. Ott részletesen leirtuk a szubjektiv eseményrendszer, a szubjektiv valószinüségeloszlás fogalmát, ez utóbbi felmérésének módjait, szóltunk a matematikai és a szubjektiv eloszlás kapcsolatáról, mértük a két eloszlás távolságát és a szubjektiv valószinüségeloszlás jóságát. Ezek ismerete az alábbiak megértéséhez nélkülözhetetlen, ezért a legalapvetőbb fogalmakat itt is ismertetjük, de nem akarunk felesleges ismétlésekbe bocsátkozni, nem készitünk egyszerü kivonatot ebből a dolgozatból, célunk a két dolgozat fogalmainak, jelöléseinek illesztése, és e dolgozat teljesebbé tétele, mindez azonban nem helyettesiti az előbb emlitett dolgozatot, hisz az alapfogalmakat ott részletesebben irtuk le. /E dolgozat különlenyomatát mellékeljük./

A valószinüségelméletben a "kisérlet" szót igen általános értelemben használják. Nézzünk néhány példát:

Két kockával dobunk egyszerre. Figyeljük a felülre került számok összegét. Ennek a kisérletnek 11 kimenetele lehet /a felső számok összege 2, 3,..., vagy 12/. A kísérlet egy kimenetelét egy eseménynek is szokták nevezni. A lehetséges események halmazát teljes eseményrendszernek nevezzük. Egy kisérlethez általában nem csak egy teljes eseményrendszert definiálhatunk. Ennél a kisérletnél például figyelhetnénk a felső számok szorzatát is, ekkor a teljes eseményrendszer 18 eseményből állna. Ezek alapján bárki tudna ujabb teljes eseményrendszert adni e kisérlethez.

Kicsit furcsának tünhet a következő példa. Rögzitsünk egy időpontot,és ekkor mérjük meg a Tisza mélységét egy szintén előre rögzitett helyen. Teljes eseményrendszernek vehetjük a Tisza lehetséges mélységét /az adott helyen/ méter pontossággal. Egy esemény például, hogy a mért eredmény 4 és 5 m. közé esik.

Valószinüségszámitási szempontból kisérlet az is, ha az ellenőr a konzervgyár raktárában találomra kiválaszt egy gulát, és megnézi, hogy abban a gulában hány konzerv nem felel meg a minőségi követelményeknek. Itt egy esemény az, hogy például 15 az elfogadhatatlan konzervek száma. Ha a gulában 1000 db konzerv volt, akkor a teljes eseményrendszer 1001 eseményből áll /a gulában 0, 1,..., vagy 1000 selejt lehet/.

Könnyen elképzelhetünk olyan kisérletet, melynek csak két kimenetele lehetséges. Elfajult esetként ide sorolhatjuk azokat a "kisérleteket" is, melyeknek egy kimenetele lehetséges.

Valójában ez utóbbi már nem véletlen esemény.

A valószinüségelmélet olyan kisérletekkel foglalkozik, amelyeknek kimenetele előre nem ismeretes, ilyenkor azt mondjuk, hogy az esemény bekövetkezése véletlen.

Legyen $\mathcal{A}$ egy teljes eseményrendszer, ennek eredményei $A_1, A_2, \ldots, A_n$. Ha a kisérletet többször elvégezzük és megszámoljuk, hogy az egyes események hányszor következtek be, és ezt elosztjuk a kisérletek számával, akkor az egyes események relativ gyakoriságát kapjuk meg. E számok 0 és 1 közé esnek, összegük pedig 1. A gyakorlat azt mutatja, hogy egy rögzitett A_i esemény relativ gyakorisága egy adott szám körül ingadozik, attól átlagosan kevésbé tér el, ha a kisérletet igen sokszor végezzük el. Ezt a számot az A_i esemény matematikai valószinüségének nevezzük. Meghatározása méréses uton a relativ gyakoriság alapján, elméleti uton pedig bizonyos feltételek elfogadása alapján történik.

Végezzük el gondolatban a Tisza mélységének mérését tiz éven keresztül ugyanazon a helyen és azonos hónapban. Tegyük fel, hogy a következő eredményeket kapjuk:

Év:	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
Mélység/m/:	5,6	4,2	4,5	5,7	3,9	4,8	3,5	4,7	6,2	5,2

Táblázatban:

mélység /m/	3-4	4-5	5-6	6-7
bekövetkezések száma	2	4	3	1
relativ gyakoriság	0,2	0,4	0,3	0,1

E mérések alapján azt mondhatjuk, hogy annak valószinüsége, hogy a Tisza az adott helyen és az adott hónapban 5 és 6 m között van 0,3. Bizonyosan ettől kissé eltérő eredményhez jutnánk, ha a mérést mondjuk 100 éven át végeznénk.

Ebből is látható, hogy egy eseményrendszerhez nem egyetemes törvény adja a matematikai valószinüségeloszlást, ezt az ember bizonyos feltételek elfogadásával határozza meg, természetesen a természeti törvények, illetve gyakorlati megfigyelései alapján. Nézzünk még egy példát ennek illusztrálására. Egy kockával dobunk, mi a valószinüsége annak, hogy a 4-es szám kerül felülre? Ilyenkor szinte minden ember a következőképpen gondolkodik: mivel minden oldalára egyenlő eséllyel eshet a kocka, ezért $\frac{1}{6}$ annak a valószinüsége, hogy a 4-est dobjuk. Ez a gondolkodás természetes és a matematikus is igy számol, de látni kell, hogy az $\frac{1}{6}$ valószinüségérték abból a feltételből adódik, hogy "minden oldalára egyenlő eséllyel eshet a kocka", amit bizonyitani nem tudunk, tulajdonképpen mi "fogtuk rá" a kisérletre, éppen a gyakorlati megfigyeléseink alapján.

Az itt felvetett probléma okozza majd az információérték relativ jellegét.

Félreértés ne essék, a fenti példákkal nem azt akartuk mondani, hogy egy esemény matematikai valószinüsége nézőpont kérdése, csupán azt hangsulyoztuk, hogy a tapaszatalat alapján mi rendeljük az adott eseményhez. Ezzel indokoltuk azt is, hogy a pszichológiában szokásos "objektiv valószinüség" kifejezés helyett /melyet a szubjektiv valószinüséggel állitanak szembe/ mi a matematikai valószinüség elnevezést használjuk. /E gondolatokat még mélyebben megérthetjük PRÉKOPA [13.] példái alapján./ Ez a gondolatsor számunkra azért volt fontos, mert hozzásegit a matematikai és a szubjektiv valószinüség elkülönitéséhez.

Egy teljes eseményrendszert $\mathcal{A}$-val, vagy $\{A_i\}$-vel, egy eseményt A_i-vel, ennek matematikai valószinüségét $P(A_i)$-vel fogjuk jelölni.

Térjünk vissza a Tisza mélységével kapcsolatos példánkhoz. Tegyük fel, hogy egy személytől megkérdezzük, hogy az adott hónapban és az adott helyen milyen mély lesz a Tisza. Pontos választ természetesen nem tud adni, csak "valószinüsiteni" tudja a mélységet. Az alapul vett /[16]/ dolgozatban leirt módon "fogadtatjuk" a v.sz.-t. Például 20 szavazási lehetőséget adunk, amit "érzése szerint" kell ugy elosztania az egyes megadott intervallumok között, hogy amit valószinübbnek tart, oda arányosan többet tegyen. Tegyük fel, hogy a következő választ kapjuk:

mélység /m/	2-nél kevesebb	2-3	3-4	4-5	5-6	6-7	7-8	8-nál több
szavazatok száma	2	4	9	2	1	1	1	0

A teljes eseményrendszer most 7 eseményből áll. /A 0 szavazatot kapott eseményeket elhagyhatjuk, ha azok matematikai valószinüsége is 0, de ha nem hagyjuk el, az sem változtat a számitás eredményén./

Mivel 20 szavazati lehetőséget osztott szét a fenti módon, ezért azt mondjuk, hogy az adott eseményekre vonatkozó szubjektiv valószinüségek rendre

$$\frac{2}{20};\quad \frac{4}{20};\quad \frac{9}{20};\quad \frac{2}{20};\quad \frac{1}{20};\quad \frac{1}{20};\quad \frac{1}{20}.$$

A szubjektiv valószinüség tehát azt fejezi ki, hogy az adott személy milyen mértékben tartja lehetségesnek, hogy a szóbanlevő esemény bekövetkezik. Ezek összege is 1.

Világos, hogy ugyanerre a kérdésre másként válaszol az, aki a Tisza adott szakasza mentén él, és másként, aki csak távolabbról ismeri a Tiszát. A szubjektiv jelző tehát jogos, hisz a kapott értékek a kiválasztott v.sz.-től függnek.

Egy A_i esemény szubjektiv valószinüségét $Q(A_i)$-val fogjuk jelölni /később e fogalomra az $R(A_i)$ jelölést is használjuk/.

Ha most visszagondolunk a Tisza mélységével kapcsolatos példánkra, azt látjuk, hogy a matematikai eseményrendszer 4 eseményből, a szubjektiv eseményrendszer viszont 7 eseményből áll. Hogyan kell ejárni, ha a kettőt össze akarjuk hasonlitani, illetve együtt akarjuk tárgyalni? Ilyen esetben a matematikai eseményrendszert ki kell bőviteni ugy, hogy az megegyezzen a szubjektiv eseményrendszerrel, csupán a hozzávett események matematikai valószinüségeit 0-nak kell venni. Együtt tárgyalva tehát a következő adatokkal kell számolni.

A_1: 2 m-nél kisebb a Tisza mélysége,

A_2 : 2 és 3 m között, A_3 : 3 és 4 m között,
A_4 : 4 és 5 m között, A_5 : 5 és 6 m között,
A_6 : 6 és 7 m között, A_7 : 7 és 8 m között van
a Tisza mélysége.

$P(A_1) = 0$ $P(A_2) = 0$ $P(A_3) = 0{,}2$ $P(A_4) = 0{,}4$

$P(A_5) = 0{,}3$ $P(A_6) = 0{,}1$ $P(A_7) = 0$

$Q(A_1) = 0{,}1$ $Q(A_2) = 0{,}2$ $Q(A_3) = 0{,}45$ $Q(A_4) = 0{,}1$

$Q(A_5) = 0{,}05$ $Q(A_6) = 0{,}05$ $Q(A_7) = 0{,}05$.

Az elnevezések és a jelölések összefoglalása:

$\mathcal{A}$ /vagy más *írott* betü/ : eseményrendszer
/Ez jelenthet matematikait, szubjektivet vagy mindkettőt./

$P(A_i)$: az A_i esemény matematikai valószinüsége

$\{P(A_i)\}$ /$i = 1, 2, \ldots, n$/ : az eseményrendszer matematikai valószinüségeloszlása

$Q(A_i)$ /vagy $R(A_i)$/ : az A_i esemény szubjektiv valószinüsége

$\{Q(A_i)\}$ az $\{A_i\}$ eseményrendszer szubjektiv valószinüségeloszlása

$P \times \mathcal{A}$ -val a matematikai valószinüségi mezőt,

$P \times Q \times \mathcal{A}$ -val a matematikai és szubjektiv valószinüségi mező együttesét jelöljük.

/ A $\times$ jel csak egy szimbólum, a matematikában nem pontosan ilyen értelemben használják, hisz ott $\mathcal{A}$ általában eseményalgebrát jelöl./

Ezzel a kiinduláshoz szükséges alapfogalmakat ismertettük, a még felmerülő problémákat /pl: valószinüségi változó/ a tárgyalás közben igyekszünk tisztázni.

II. Az információ értéke adott szubjektumra és rögzitett $\mathcal{A}$ eseményrendszerre vonatkozóan

Induljunk ki egy egyszerü elképzelt pszichológiai kisérletből, ahol fiktiv adatokkal dolgozunk. /Az elméleti meggondolásoknál ugy sincs jelentősége a konkrét eredményeknek, sőt az adatok tetszés szerinti változtatása még nagyobb lehetőséget nyujt a különböző esetek bemutatására./

A már emlitett dolgozatunkban leirtuk a következő /valóban végrehajtott/ kisérletet: Egy körlapra négy egyenlő vastagságu körgyürüt rajzoltunk és ezt céltáblának neveztük. Azt mondtuk a kisérleti személynek, hogy több találomra leadott lövésből 100 eltalálta a céltáblát /ami esetleg le is volt takarva/, becsüljük meg a 100 lövés közül hány esett az egyes körgyürükbe. A mérések és számitások eredményei lényegében a következő kérdésre adják meg a választ:
Ha találomra lövünk egy céltábla felé és véletlenül eltaláljuk, mi a valószinüsége, hogy az egyes körgyürükbe esik a találat.

Az eseményrendszer itt a következő:

A_1 : a belső körbe esik a találat

:

A_4 : a legkülső körgyürübe esik a találat

A matematikai valószinüségeket az alapján számitottuk, hogy feltettük, hogy egy esemény valószinüsége arányos a neki megfelelő terület nagyságával, vagyis az 1-et a területek arányában osztottuk fel.

A szubjektiv valószinüségek meghatározását [16.]-ban részletesen leirtuk.

A következő eredményeket kaptuk:

i	1	2	3	4
$P(A_i)$	0,062	0,187	0,313	0,438
$Q(A_i)$	0,138	0,242	0,342	0,279

A szubjektiv valószinüségeloszlást ugyan több k.sz. által megadott adatok alapján számitottuk, de a továbbiakban tételezzük fel, hogy ez egy szubjektum véleménye. A feltételezett vizsgálati személy a legkisebb valószinüséget az A_1 eseménynek, legnagyobbat az A_3 eseménynek tulajdonitja.

Képzeljük el, hogy valóban végrehajtjuk a kisérletet, és a kisérlet eredményét közöljük a kisérleti személlyel /most már csak 1 véletlen találatról van szó/. Mindenképpen bizonyos meglepetést okozunk a k.sz.-nek, hisz nem tudta előre, melyik körgyürübe fog esni a találat. A meglepetés mértéke attól függ, hogy valójában melyik esemény következett be. Nyilvánvaló, hogy az A_1 esemény bekövetkeztének közlése nagyobb meglepetést fog okozni, mint az A_3 esemény bekövetkeztének közlése, hisz az A_3-at "jobban várta". Még jobban érthető ez a gondolat, ha olyan példára gondolunk, ahol az egyik esemény szubjektiv valószinüsége nagyon kicsi, a másiké pedig nagy. Egyszerü példa erre, ha a k.sz.-nek azt mondjuk, hogy egy adott dobozban egér van, mondja meg milyen szinü. Bizonyosan nagyon kicsi a fehér szin szubjektiv valószinüsége, a szürke sziné pedig nagy. Igy alig okozunk meglepetést, ha szürke egeret mutatunk fel, viszont nagy meglepetést okozunk, ha fehér egeret veszünk ki a dobozból.

A felsorolt konkrét példákkal párhuzamosan irjuk le és vizsgáljuk meg e kérdést általánosan is.

Vegyünk egy tetszőleges $\mathcal{A}$ eseményrendszert, aminek adott, illetve meghatároztuk a matematikai és a szubjektiv valószinüségeloszlását. Hajtsuk végre a kisérletet, a kisérlet eredményét közöljük a k.sz.-lyel. Tegyük fel, hogy az A_i esemény következett be, aminek szubjektiv valószinüsége $Q\ (A_i)$. A példákból láttuk, ha $Q\ (A_i)$ kicsi, akkor nagy meglepetést okozunk a k.sz.-nek, és ez forditva is fennáll. A meglepetés és a szubjektiv valószinüség tehát "forditottan arányos". /Itt a "forditottan arányos" kifejezés hétköznapi értelemben szerepel, hisz a meglepetésnek még nem adtunk értéket./ Másként a meglepetés a szubjektiv valószinüség reciprokával "arányos". Mivel a $\log_2 x$ függvény monoton növő, ezért az eddigi megállapitásainkkal összhangban maradunk, ha a meglepetés értékét

$$\log_2 \frac{1}{Q\ (A_i)} \text{ - vel definiáljuk.}$$

Összefoglalva azt mondjuk, hogy

$$\log_2 \frac{1}{Q(A_i)} \text{ meglepetést okozunk a k.sz.-nek,}$$

ha egy $Q(A_i)$ szubjektiv valószinüségü A_i esemény bekövetkeztét közöljük.

Felvetődhet a kérdés, hogy miért éppen a fenti definiciót adjuk meg, hisz magával az $\frac{1}{Q(A_i)}$ -vel is definiálhatnánk a meglepetés értékét, sőt $\log_2 x$ függvény helyett bármely szigoruan monoton növekvő függvényt is használhatnánk. /A szóbajöhető függvényről azért kell kikötnünk a monoton növekedést, hogy kisebb $Q(A_i)$ -hez biztosan nagyobb meglepetés érték tartozzon. /A $\log_2 x$ függvény használatának előnyét az információelmélet indokolja, ez a függvény tesz eleget bizonyos alapvető követelményeknek. A 2-es alapu logaritmus helyett más alapu logaritmust is használhatnánk, de ez a megszokott, igy az információ értékét majd bit-ben kapjuk. Mivel mindig a 2-es alapu logaritmussal fogunk dolgozni, e dolgozatban ezután elhagyjuk az alap értékének jelölését, tehát $\log x = \log_2 x$ jelölést használjuk. Mivel azonban a forgalomban levő táblázatok a 10-es alapu logaritmus értékét adják meg, ezért közöljük az átszámitás módját:

$$\log x = \log_2 x = \frac{\lg x}{\lg 2} \doteq \frac{\lg x}{0{,}301} = 3{,}322 \cdot \lg x \,,$$

tehát

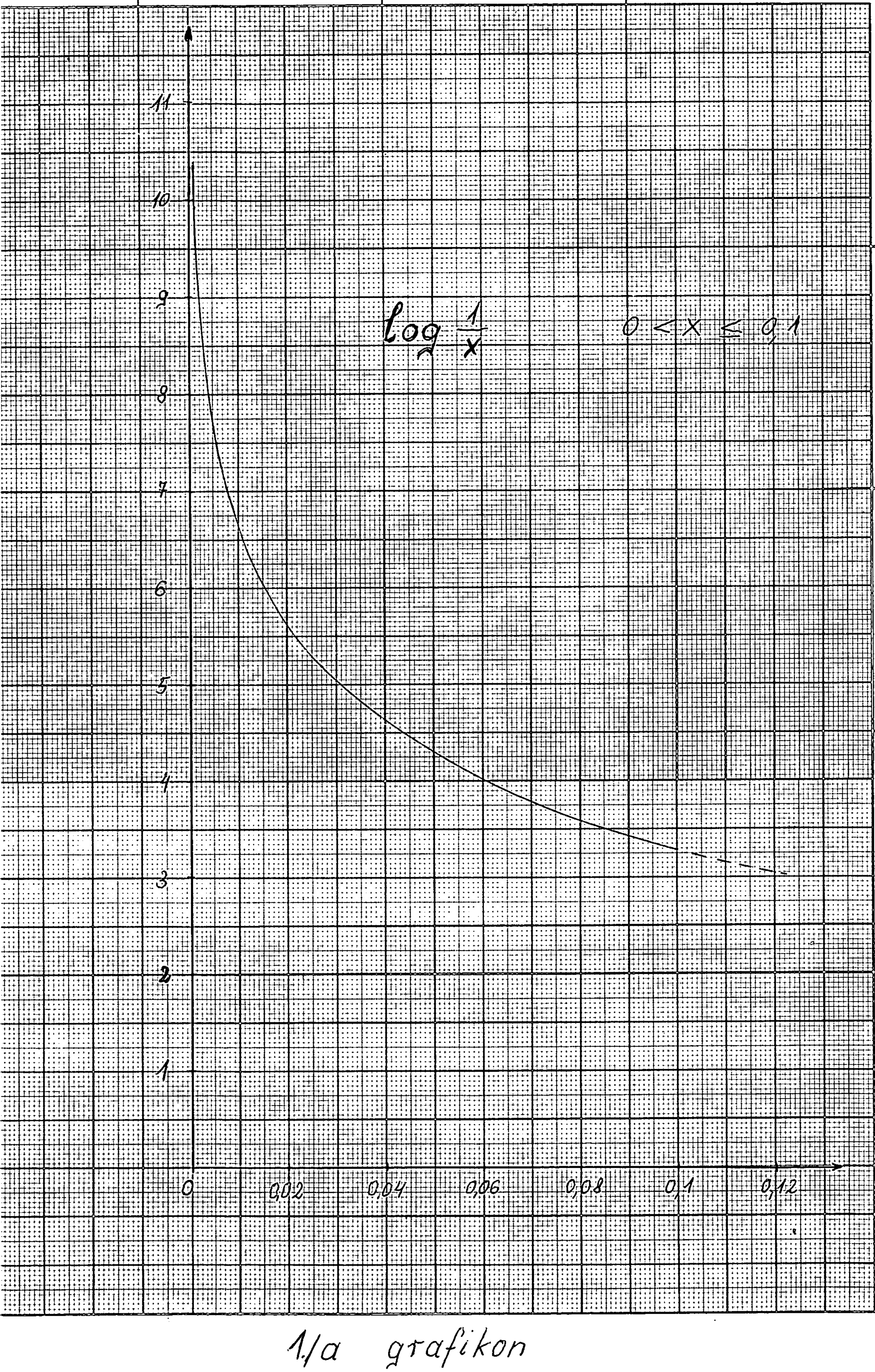
$$\log x = 3{,}322 \lg x \,, \quad \text{ahol lg a 10-es alapu}$$

logaritmust jelöli.

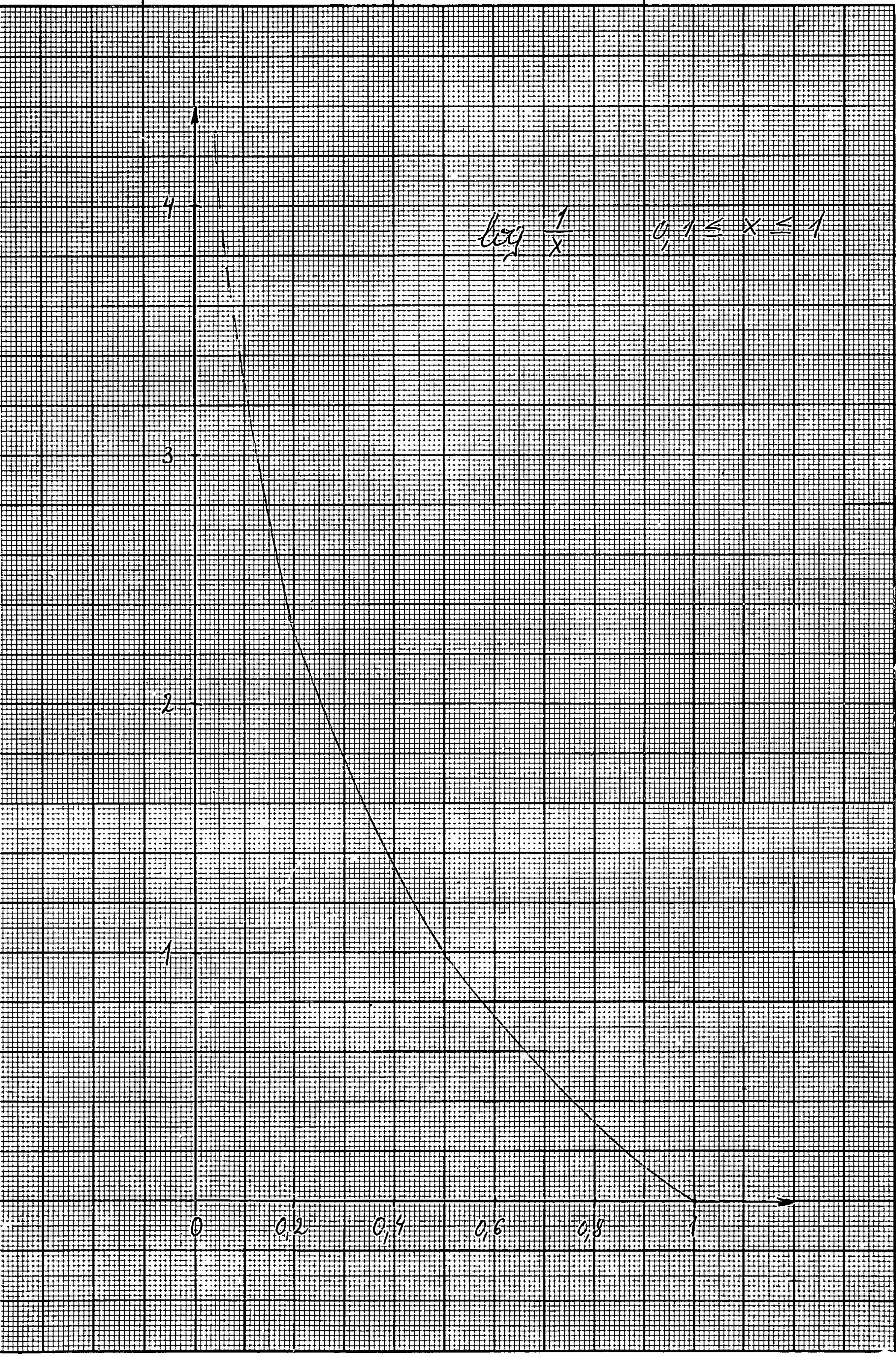
A számitások követésének megkönnyitése céljából az 1/a és 1/b grafikonokkal ábrázoltuk a $\log \frac{1}{x}$ függvényt a $(0; 1]$ intervallumon.

A meglepetés fogalmával kapcsolatban még egy problémára ki kell térnünk. Ha egy A_i esemény szubjektiv valószinüsége 0, akkor a meglepetés értéke $\log \frac{1}{0}$ lenne, aminek nincs értelme, ilyenkor azt mondhatjuk, hogy a meglepetés végtelen nagy, nincs értéke /matematikailag $\lim\limits_{x \to 0} \log \frac{1}{x} = \infty$ / Ez pszichológiai szemszögből nézve sem ellentmondásos, mert ha $Q(A_i) = 0$, ez azt jelenti, hogy a szubjektum lehetetlennek tartja az A_i esemény bekövetkeztét. Ha mégis ennek bekövetkeztéről értesitjük, a meglepetést valóban nem mérhetjük. A gyakorlatban ez annyiban okoz problémát, hogy vigyázni kell a felmérésnél. Amikor egy ember azt mondja, hogy "lehetetlen, hogy 5-ös találatom lesz ezen a lottón", valójában a szubjektiv valószinüség nem 0, hisz tudja, hogy ezeket a számokat is kihuzhatják. Tehát ha egy felmérés eredményeként $Q(A_i) = 0$ jön ki, de $P(A_i) \neq 0$, akkor meg kell vizsgálni, hogy ez a felmérés hibája vagy valóságos véleményt fejez e ki.

Az itt bevezetett $\log \frac{1}{Q(A_i)}$ formula az információelméletben alapvető, alkalmazása a pszichológiában is elterjedt, de /legalábbis az általunk ismert irodalomban/ a matematikai valószinüséget helyettesitik a képletbe. Azzal, hogy mi a szubjektiv valószinüséggel számolunk /ez tulajdonképpen az egyik alapgondolatunk/ lényegesen megváltozik az elméleti felépités és a gyakorlati felmérés.



1./a grafikon



1. /b. grafikon

E szemléletmód változtatásra gondoltunk, amikor azt irtuk, hogy "igy az információelmélet adekváltabb módszerévé válik e kérdéskör vizsgálatának". Valóban, ha egy szubjektum valamely tulajdonságát akarjuk mérni, akkor az adott szubjektum adatait /és nem más, általunk kispekulált adatokat/ kell felhasználni, azokat kell elemezni. Ha a legelső, céltáblás kisérletre gondolunk, nyilvánvaló, hogy a $\log \frac{1}{0,062}$ /ahol $0,062 = P(A_1)$ /semmilyen kapcsolatban sincs az adott szubjektummal, bárkivel végezzük a kisérletet, a matematikai valószinüség nem változik.
Más a helyzet a $\log \frac{1}{0,138}$ -al, ahol $0,138 = Q(A_1)$, ez a szubjektumra jellemző adat, annak bizonyos ismeretétől, tájékozottságától, gondolkodásmódjától függ.

Térjünk vissza az $\mathcal{A}$ eseményrendszerre. A kisérletet végrehajtva közöltük a k.sz.-lyel, hogy az A_1 esemény következett be. Ekkor azt mondtuk, hogy a k.sz.-t $\log \frac{1}{Q(A_1)}$ meglepetés érte.

Ebből arra következtethetnénk, hogy a k.sz. annál tájékozottabb az $\mathcal{A}$ eseményrendszerre vonatkozóan /annál helyesebb az elképzelése/, minél kisebb a $\log \frac{1}{Q(A_1)}$ értéke. Ellenpéldaként vizsgáljunk meg egy fiktiv kisérletet.

Tegyük fel, hogy a kérdés két /A_1 és A_2/ eseményből álló eseményrendszerre vonatkozik.
Legyen $P(A_1) = 0,2$ és $P(A_2) = 0,8$.

Tegyük fel továbbá, hogy két k.sz. informáltságát kivánjuk felmérni az adott eseményrendszerre vonatkozóan. Ha az első k.sz. esetében $Q'(A_1) = 0,1$, $Q'(A_2) = 0,9$ szubjektiv eloszlást, a másodiknál pedig $Q''(A_1) = 0,7$, $Q''(A_2) = 0,3$ szubjektiv valószinüségeket mérünk fel, akkor természetesen az elsőt kell jobban informáltnak tekintenünk, ő itéli meg reálisabban a helyzetet. A kisérlet végrehajtásánál bármely esemény bekövetkezhet, hisz egyik esemény matematikai valószinüsége sem 0, pontosabban egyik esemény sem lehetetlen /az utóbbi azért helyesebb kifejezés, mert előfordulhat, hogy a 0 matematikai valószinüségü esemény is bekövetkezik/. Ha például az A_1 esemény következik be, akkor ez az első k.sz. esetében $\log\frac{1}{0,1} = 3,322$ meglepetést, a második k.sz. esetében pedig csak $\log\frac{1}{0,7} = 0,515$ meglepetést okoz. Ugy tünhet, hogy itt ellentmondás van, hisz az első k.sz.-t tájékozottabbnak fogadtuk el, és mégis őt érte nagyobb meglepetés a kisérlet végrehajtásakor. Ez a látszólagos ellentmondás akkor oldódik fel, ha a kisérlet többszöri végrehajtására, illetve ezek eredményeinek közlésére gondolunk. Az előbbi eset, hogy az A_1 esemény bekövetkeztét kell közölnünk jóval kevesebbszer fordul elő, mivel ennek matematikai valószinüsége kisebb, tehát átlagosan valóban az első k.sz.-nek okozunk kisebb meglepetést, ha a kisérletek kimeneteleit közöljük.

Fontos tanulság, hogy valójában a k.sz. tájékozottságát nem egy konkrét meglepetés értéke alapján kell mérni, hanem a meglepetés várható értékét kell figyelembe venni.

Maradjunk az előbbi példánál. Az első k.sz.-t $\log \frac{1}{0,1}$ vagy $\log \frac{1}{0,9}$ meglepetés érheti, attól függően, hogy melyik esemény bekövetkeztét közöljük vele. Ugyanez a második k.sz. esetében $\log \frac{1}{0,7}$ vagy $\log \frac{1}{0,3}$ lehet. Vegyük e meglepetésértékek sulyozott átlagát, ahol a megfelelő sulyokat a matematikai valószinüségek adják:

Első esetben $\quad 0,2 \cdot \log \frac{1}{0,1} + 0,8 \cdot \log \frac{1}{0,9} = 0,786$

Második esetben $\quad 0,2 \cdot \log \frac{1}{0,7} + 0,8 \cdot \log \frac{1}{0,3} = 1,493$

/Az átlagképzésnél mindkét esetben 1-el kellett volna osztanunk, ezért nem irtuk ki az osztást./

Ezeket a sulyozott átlagokat a valószinüségszámitásban várható értékeknek nevezik.

Elmondhatjuk tehát, hogy az első k.sz. meglepetésének várható értéke /0,786/ jóval kisebb, mint a másodiké /1,493/, tehát matematikailag is indokoltuk, hogy az első k.sz.-t kell tájékozottabbnak tekintenünk az adott eseményrendszerre vonatkozóan.

Általában egy $\mathcal{A}$ eseményrendszerre vonatkozóan a k.sz. meglepetésének várható értékét a

$$P(A_1) \cdot \log \frac{1}{Q(A_1)} + P(A_2) \cdot \log \frac{1}{Q(A_2)} + \dots + P(A_n) \cdot \log \frac{1}{Q(A_n)}$$

összeggel számitjuk ki, amit rövidebben a

$$\sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)}$$ formulával szoktunk kifejezni.

A céltáblás kisérletnél tehát a vizsgált személy meglepetésének várható értéke

$$0{,}062 \cdot \log \frac{1}{0{,}138} + 0{,}187 \cdot \log \frac{1}{0{,}242} + 0{,}313 \cdot \log \frac{1}{0{,}343} +$$

$$+ \log \frac{1}{0{,}279} = 0{,}177 + 0{,}383 + 0{,}483 + 0{,}807 = 1{,}85$$

Az információelméletben egy $\mathcal{A}$ eseményrendszer /melynek matematikai valószinüségeloszlása $\{P(A_i)\}$, ahol $i=1, 2, \dots, n$/ entrópiáját a

$$\sum_{i=1}^{n} P(A_i) \cdot \log \frac{1}{P(A_i)}$$ összeggel definiálják.

Ez az általunk bevezetett fogalommal rokon fogalom, a rendszer bizonytalanságát méri. Sőt, ha a szubjektiv eloszlás megegyezik a matematikaival, akkor e két mennyiség is megegyezik, tehát az eseményrendszer entrópiája egyben a meglepetés várható értéke.

Hogy az analógia az elnevezésekben is megmutatkozzék, a meglepetés várható értéke elnevezés helyett helyesebb, ha a

$$\sum_{i=1}^{n} P\,(A_i)\log\frac{1}{Q\,(A_i)} \quad \text{mennyiséget a}$$

k.sz. $\mathcal{A}$ eseményrendszerre vonatkozó szubjektiv entrópiájának nevezzük.

Bár bizonyos esetben szükség lehet a meglepetés konkrét $\log\frac{1}{Q\,(A_i)}$ értékére, az elméleti meggondolásoknál és a gyakorlati felméréseknél általában a szubjektiv entrópia játszik fontos szerepet, ezért ezt még részletesen vizsgáljuk.

A gyakorlatban sokszor előfordul, hogy a matematikai eloszlás elfajult, ami azt jelenti, hogy valamely esemény matematikai valószinüsége 1, a többié 0. A feleletválasztásos tesztek például ilyenek. Tegyük fel például, hogy KRESZ vizsgán a következő kérdést teszik fel: Mit jelent az ábrán látható tábla /csucsára állitott, fehér alapon piros szegélyü háromszög/? A következő válaszok közül kell a helyeset megjelölni:

1. Állj, elsőbbségadás kötelező / A_1 /
2. Elsőbbségadás kötelező / A_2 /
3. Általános veszély / A_3 /

A matematikai eloszlás: $P\,(A_1) = 0$, $P\,(A_2) = 1$ és $P\,(A_3) = 0$. Ebben az esetben indokolt, hogy a vizsgált személy nem valószinüsithet, helyesebben 1 szubjektiv valószinüségünek kell lenni az A_2 eseménynek /egy vezetőnek pontosan kell ismernie a KRESZ-t/, de más hasonló tipusu felméréseknél helyes lenne, ha a szubjektiv eloszlást mérnénk fel és nem csak az egyik válasz megjelölését engednénk meg, mert ugy pontosabb képet kapnánk a v.sz. véleményéről. Most tételezzük fel /a valósággal ellentétben/, hogy a fenti kérdésre a vizsgázó a szubjektiv valószinüségeket adhatja meg. Ebben az esetben a szubjektiv entrópia megegyezik az A_2-höz tartozó meglepetés értékével, mert

$$0 \cdot \log \frac{1}{Q(A_1)} + 1 \cdot \log \frac{1}{Q(A_2)} + 0 \cdot \log \frac{1}{Q(A_3)} = \log \frac{1}{Q(A_2)}$$

Állapodjunk meg abban, hogy

$0 \cdot \log \frac{1}{0}$ értékét 0 -nak vesszük.

Ezt azért kell külön kezelnünk, mert matematikailag ennek a kifejezésnek nincs értéke.

Megállapodásunkat viszont a

$$\lim_{x \to 0} x \log \frac{1}{x} = 0$$

elfogadhatóvá teszi.

Megjegyezzük, hogy e megállapodásunk csak a diszkrét esetre vonatkozik.

A feleletválasztásos pedagógiai felméréseknél is csak egy választ szoktak elfogadni. Itt határozottan helyesebb lenne az adott válaszokra, mint eseményrendszerre vonatkozó szubjektiv valószinüségeloszlás felmérése. Ezzel a találgatást kiküszöbölnénk, ugyanis a megszokott módszer esetében az a tanuló, aki egyáltalán nem ismeri a helyes választ, találomra jelöl meg egyet, így el is találhatja a helyeset.

/Bár a kiértékelésnél figyelembe veszik annak valószinüségét, hogy a készületlen tanuló bizonyos számu választ el is találhat./ Ha a szubjektiv valószinüségeloszlást mérnénk fel, akkor $\log \frac{1}{Q(A)}$ alapján következtethetnénk arra, hogy a tanuló erre a kérdésre milyen helyesen tud válaszolni /"A" itt a helyes választ jelöli/, mégpedig annál jobb a válasza, minél kisebb a $\log \frac{1}{Q(A)}$.

Vizsgáljuk meg a szubjektiv entrópia nagyságát. Itt tulajdonképpen csak azt a speciális esetet vizsgáljuk, amikor $\mathcal{A}$ két eseményből áll. Mivel grafikus módszert alkalmazunk, ideiglenesen célszerü áttérni a függvénytanban szokásos jelölésekre. Legyen $P(A_1) = x$, $Q(A_1) = y$, ekkor $P(A_2) = 1 - x$ és $Q(A_2) = 1 - y$ /azt használtuk ki, hogy a valószinüségeloszlás összege 1/.

A bevezetett jelölésekkel a szubjektiv entrópia a következő formában irható fel:

$$z = x \cdot \log \frac{1}{y} + /1 - x/ \cdot \log \frac{1}{1-y} .$$

Ez egy kétváltozós függvény, ami a háromdimenziós térben egy felületet ad meg. Mi e felület sikmetszeteit tudjuk ábrázolni a sikban ugy, hogy x értékét rögzitjük, és a kapott /egyváltozós/ függvényt ábrázoljuk derékszögü koordinátarendszerben. A végtelen sok lehetőség közül csak 5-öt választunk ki, de ezekből már leolvashatunk néhány fontos tulajdonságot.

Legyen először x = 0,1 , ekkor

$$z = 0{,}1 \log \frac{1}{y} + 0{,}9 \log \frac{1}{1-y} \,.$$

Először értéktáblázatot készitünk, vagyis y helyére különböző számokat helyettesitve kiszámitjuk az összeg értékét.

Kapott eredmények: / x = 0,1 /

y	0,01	0,05	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,99
z	0,677	0,499	0,469	0,522	0,637	0,795	1	1,264	1,614	2,12	3,0	3,9	5,98

E görbét kék, folyamatos vonallal ábrázoltuk /2.grafikon/.

Ugyanigy jártunk el az x = 0,3 /zöld folytonos vonal/ és az x = 0,5 /fekete vonal/ esetében is.
Az x = 0,7 és x = 0,9 esetekben felesleges lett volna uj értéktáblázatokat készitetünk, mert az x = 0,7-nek megfelelő függvény esetében a 0,6-nél annyit kapunk, mint az x = o,3-nek megfelelő függvény esetében a 0,4-nél. E két függvény tehát tükörképe egymásnak. Ezt a tényt a grafikonon a megfelelő szinek alkalmazásával igyekeztünk szemléltetni.

A görbék legmélyebb pontjait, a függvények minimumát piros szinnel jelöltük, amelyek /ha a fel nem rajzolt görbéket is figyelembe vesszük/ egy parabolához hasonló görbét adnak.

Könnyen észrevehető, hogy minden függvény olyan y érték mellett veszi fel minimumát, amennyi az x volt. /Az x=0,3-hez tartozó görbe az y=0,3-nél a legalacsonyabb/. Ez a tény az eredeti jelölések mellett azt jelenti, hogy a

$$\sum_{i=1}^{2} P(A_i) \log \frac{1}{Q(A_i)} \quad \text{akkor a legkisebb,}$$

ha $Q(A_i) = P(A_i)$ /i = 1,2/ , másként ha megegyezik a matematikai entrópiával.

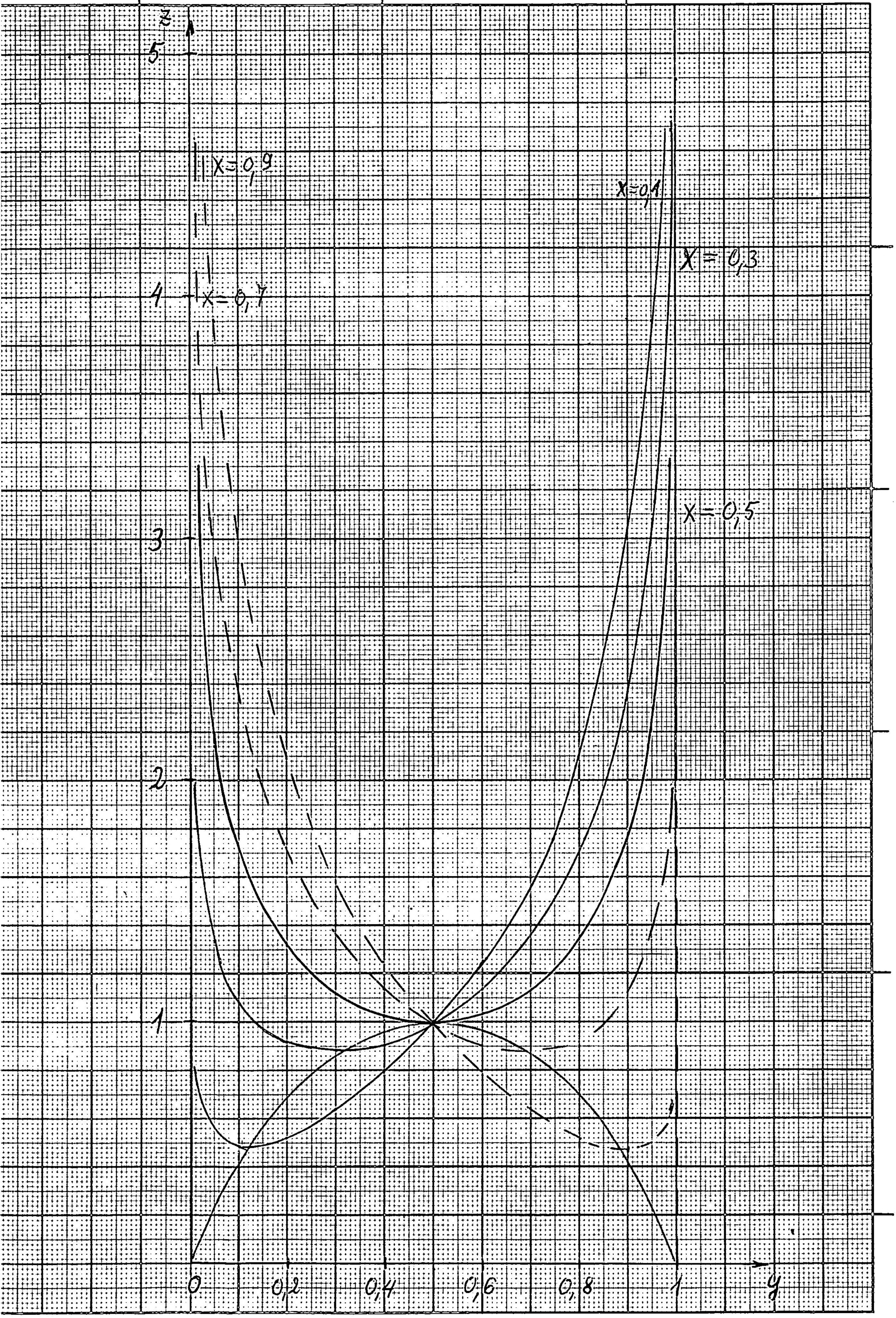
Ez a megállapitás egy általános tétel speciális esete. Igazolható / ld. [5.] 714. old./, hogy

$$\sum_{i=1}^{n} P(A_i) \log \frac{1}{P(A_i)} \leq \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} \quad , \text{ az}$$

egyenlőség csak akkor áll fenn, ha $Q(A_i) = P(A_i)$ minden i-re.

Lényeges következmény számunkra, hogy a szubjektiv entrópia akkor a legkisebb /adott $\mathcal{A}$ -ra vonatkozóan/, ha a szubjektiv valószinüségeloszlás megegyezik a matematikaival.

Az információelméletben bizonyitott tény, hogy egy rendszer /matematikai/ entrópiája nem lehet nagyobb log n-nél, ezt a maximális értéket akkor veszi fel, ha minden esemény egyenlő valószinüségü. A grafikonról viszont leolvasható, hogy a szubjektiv entrópiának nincs felső határa. Itt is látható tehát e két entrópia közötti lényeges különbség. A szubjektiv entrópia annál nagyobb, minél nagyobb a szubjektiv és matematikai eloszlás távolsága.



2. grafikon

Mivel azonban /az x = 0,5 -et kivéve/ a görbék nem szimmetrikusak, nem mindegy, hogy a szubjektiv eloszlás merre felé tolódik el a matematikaitól. E kérdés vizsgálatára még visszatérünk.

E kitérő után térjünk rá az alcimben kitüzött feladatunk megoldására.

Tekintsünk vissza a céltáblás kisérletre. Tételezzük fel, hogy az első felmérés után megmondjuk a v.sz.-nek, hogy a találatok körülbelül egyenletesen oszlanak meg a céltáblán. Feltételezhető, hogy ezzel a közléssel arra serkentettük a v.sz.-t, hogy a területek arányában ossza el szavazási lehetőségeit. A szubjektiv eloszlás az ujabb felmérés után mondjuk a következőképpen változott:

körgyürük	1	2	3	4
szubjektiv valószinüség	0,11	0,2	0,31	0,38

Számitsuk ki ezekkel az adatokkal is a szubjektiv entrópiát:

$$0,062 \cdot \log \frac{1}{0,11} + 0,187 \cdot \log \frac{1}{0,2} + 0,313 \cdot \log \frac{1}{0,31} +$$

$$+ 0,438 \cdot \log \frac{1}{0,38} = 1,77 \ .$$

Mivel a szubjektiv valószinüségeloszlás közelebb került a matematikaihoz, ezért a szubjektiv entrópia csökkent. A csökkenés értéke: $1,85 - 1,77 = 0,08$, Legyen adott egy $\mathcal{A}$ eseményrendszer $\{P(A_i)\}$ matematikai valószinüségeloszlással, tehát a $P \times \mathcal{A}$ mező. Adott továbbá egy szubjektum, akiben az adott eseményrendszerrel kapcsolatban egy $\{Q(A_i)\}$ szubjektiv valószinüségeloszlás lép fel.

Azt mondjuk, hogy a szubjektum információt vesz fel az $\mathcal{A}$-ra vonatkozóan, ha benne az $\mathcal{A}$-ra vonatkozó szubjektiv valószinüségeloszlás megváltozik.

Ezzel egzakt módon definiáltuk az információ fogalmát. Az egzaktság abban van, hogy felméréssel egyértelmüen eldönthető, kapott e a szubjektum információt.

Itt még csak speciális esetben /$\mathcal{A}$ -ra vonatkozóan/ definiáltuk az információ fogalmát, de ezt a fogalmat már könnyü lesz általánositani.

Lényeges, hogy az információra semmiféle megkötést nem adtunk, létrejöhet külső inger hatására, de a szubjektum gondolkodása is előidézheti.

/Ez tehát lényegesen általánosabb definició, mint ami a matematikában használatos, ahol információnak egy esemény bekövetkeztének közlését tekintik./

Az információt I-vel fogjuk jelölni.

Az eredeti szubjektiv valószinüségeloszlást a priori szubjektiv eloszlásnak, az ujat pospriori szubjektiv eloszlásnak is nevezhetjük.

Számunkra másik fontos fogalom az információ értéke, amit az eredeti és az uj szubjektiv entrópia különbségeként definiálunk.

Könnyen érthető a következő szimbólum:

$$P \times Q \times \mathcal{A} \xrightarrow{\quad I \quad} P \times R \times \mathcal{A}$$

A matematikai valószinüségeloszlás tehát változatlan, a szubjektiv eloszlás viszont

$\{Q(A_i)\}$ -ről $\{R(A_i)\}$ -re változott, ami az I információ közlését; létrejöttét jelenti.

A szubjektiv valószinüségeloszlást megváltoztató I információ $\mathcal{A}$ -ra vonatkozó értékét $\phi_{\mathcal{A}}(I)$ -vel jelöljük. Az előbbi definició szerint tehát

$$\phi_{\mathcal{A}}(I) = \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} - \sum_{i=1}^{n} P(A_i) \log \frac{1}{R(A_i)} ,$$

kiemelés után

$$\phi_{\mathcal{A}}(I) = \sum_{i=1}^{n} P(A_i) \log \frac{R(A_i)}{Q(A_i)}$$

Már ebből is látszik, amit a bevezetőben emlitettünk, hogy az információ nálunk általánosabb megfogalmazást kap, mint az a definició, hogy "az információ bizonytalanságot feloldó közlést jelent". Mivel a szubjektiv entrópia a szubjektum bizonytalanságát méri, ezért I értékét a bizonytalanság megváltozásának értékeként definiáltuk.

Mi tehát már a bizonytalanság megváltozását is információnak tekintjük, aminek speciális esete a bizonytalanság megszünése. Sőt mi azt sem mondtuk, hogy a szubjektiv entrópiának csökkennie kell. Látni fogunk olyan példát, ahol a szubjektiv entrópia növekszik, igy I értéke negativ szám lesz.

A szétszórtan leirt céltáblás kisérletet a fenti definiciók alapján a következőképpen foglalhatjuk össze:

Eredeti állapot:

i	1	2	3	4
$P(A_i)$	0,062	0,187	0,313	0,438
$Q(A_i)$	0,138	0,242	0,343	0,279

Közlés: "a találatok egyenletesen oszlanak meg a céltáblán"

$R(A_i)$	0,11	0,2	0,31	0,38

A közlés tehát információt hordoz, ennek értéke

$$\phi_{\mathcal{A}}(I) = 0{,}062 \cdot \log \frac{0{,}11}{0{,}138} + 0{,}187 \cdot \log \frac{0{,}2}{0{,}242} +$$

$$+ 0{,}313 \cdot \log \frac{0{,}31}{0{,}342} + 0{,}438 \cdot \log \frac{0{,}38}{0{,}279} = 0{,}08 \ .$$

Csak megjegyezzük, hogy ha az egyik esemény matematikai valószinüsége 1, a többié pedig 0, akkor

$$\phi_{\mathcal{A}}(I) = \log \frac{R(A_i)}{Q(A_i)} \ , \quad \text{ahol} \quad P(A_i) = 1 \ .$$

Két speciális esetre irányitjuk a figyelmet.

Kisérletileg igazolt tény / [10.] 168. old./, hogy "a szubjektum, amikor valamilyen ismeretlen feladat megoldásához fog, az összes lehetséges események egyenlő valószinüségéből indul ki." /Általában problémát okozhat az "összes lehetséges esemény" kifejezés fogalma, ennek meghatározása. Nálunk még ez nem okoz gondot, hisz adott $\mathcal{A}$-ra vonatkozó információfelvételt vizsgálunk./ Joggal feltételezhetjük tehát, hogy a teljesen tájékozatlan ember esetében $\frac{1}{n}$ a kezdeti szubjektiv valószinüség minden A_i esetében. Tehát a kezdeti szubjektiv entrópia

$$\sum_{i=1}^{n} P(A_i) \log \frac{1}{\frac{1}{n}} = \log n.$$

Amikor tehát mi egy $\{Q(A_i)\}$ szubjektiv valószinüségeloszlást mérünk fel, akkor ez azt jelenti, hogy a szubjektum az $\mathcal{A}$ -ra vonatkozóan már vett fel információt, mégpedig

$$\log n - \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} = \sum_{i=1}^{n} P(A_i) \log n \cdot Q(A_i)$$

értéküt. Ezt a mennyiséget a szubjektiv eloszlásban tartalmazott információ értéknek is nevezhetjük.

Kihangsulyozzuk, hogy ez csak az egyenletes szubjektiv eloszlásról a jelenlegi szubjektiv eloszlásra való átmenethez tartozó információ értékét adja. Valójában, hogy ez a szubjektiv eloszlás kialakuljon, nagyobb információra van szükség, mert már a lehetséges eseményrendszer felismerése is információfelvétel eredménye, aminek értékét itt nem vettük számitásba."

Előfordulhat, hogy I hatására $R(A_i) = P(A_i)$ minden i-re, vagyis a posapriori szubjektiv eloszlás megegyezik a matematikaival. Ekkor

$$\not{O}_{\mathcal{A}}(I) = \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} - \sum_{i=1}^{n} P(A_i) \log \frac{1}{P(A_i)} .$$

Az előbbiekben viszont megállapitottuk, hogy a szubjektiv entrópia akkor a legkisebb, ha megegyezik a matematikai entrópiával. Ilyen feltétel mellett vonunk tehát ki legkevesebbet az eredeti szubjektiv entrópiából, ekkor kapunk legnagyobb eredményt. Tehát /és ez várható is volt/ akkor közöljük a legnagyobb információt a szubjektummal az adott $\mathcal{A}$-ra vonatkozóan, ha abban a matematikai eloszlás alakul ki.

E gondolatsornál feltettük, hogy a matematikai eloszlás nem változik I közlésével. Később elő fog fordulni, hogy a matematikai eloszlás is megváltozik, de a fenti kijelentés akkor is igaz marad, csak megfelelően kell értelmezni.

Ezzel a legspeciálisabb esetben definiáltuk egy I információ értékét adott szubjektumra és rögzitett $\mathcal{A}$ eseményrendszerre vonatkozóan. A specialitás abban van, hogy eddig hallgatólagosan feltételeztük a matematikai eloszlás változatlanságát és a valószinüségi változóra nem tértünk még ki, ami egyes esetekben általánosabb leirásra ad alkalmat. Mielőtt azonban továbbmennénk, néhány megjegyzést kell tennünk e fogalommal kapcsolatban.

A szövegben mindig ~~i~~hangsulyoztuk, és a jelölés is arra utal, hogy ez az információhoz rendelt érték függ a szubjektumtól és az eseményrendszertől. Ennek oka nyilvánvaló. Két különböző személy esetén egy adott közlés nem biztos, hogy ugyanazt a hatást fejti ki ugyanazon eseményrendszerre vonatkozóan, hisz már az is ritka eset, ha az a priori szubjektiv valószinüségeloszlás megegyezik. Az is nyilvánvaló, hogy egy adott információ nem mindig csak egy eseményrendszer szubjektiv eloszlását változtatja meg. Összefoglalva, a fenti definició nem egyértelmüen rendel számot egy közléshez.

E definiciónak érdekes tulajdonsága, hogy egy információhoz negativ érték is tartozhat, tehát az is előfordulhat, hogy a szubjektum negativ értékü információt vesz fel. Ez a tény a pszichológia szempontjából kimondottan hasznos, hisz igy a "helytelen" információ is mérhető. Ezt egy példán megmutatjuk.

$\mathcal{A}$:	A_1	A_2
$P(A_i)$	0,2	0,8
$Q(A_i)$	0,4	0,6
$\big\downarrow I$		
$R(A_i)$	0,9	0,1

A felvett információ értéke:

$$\phi_{\mathcal{A}}(I) = 0{,}2 \cdot \log \frac{0{,}9}{0{,}4} + 0{,}8 \log \frac{0{,}1}{0{,}6} = -1{,}834 .$$

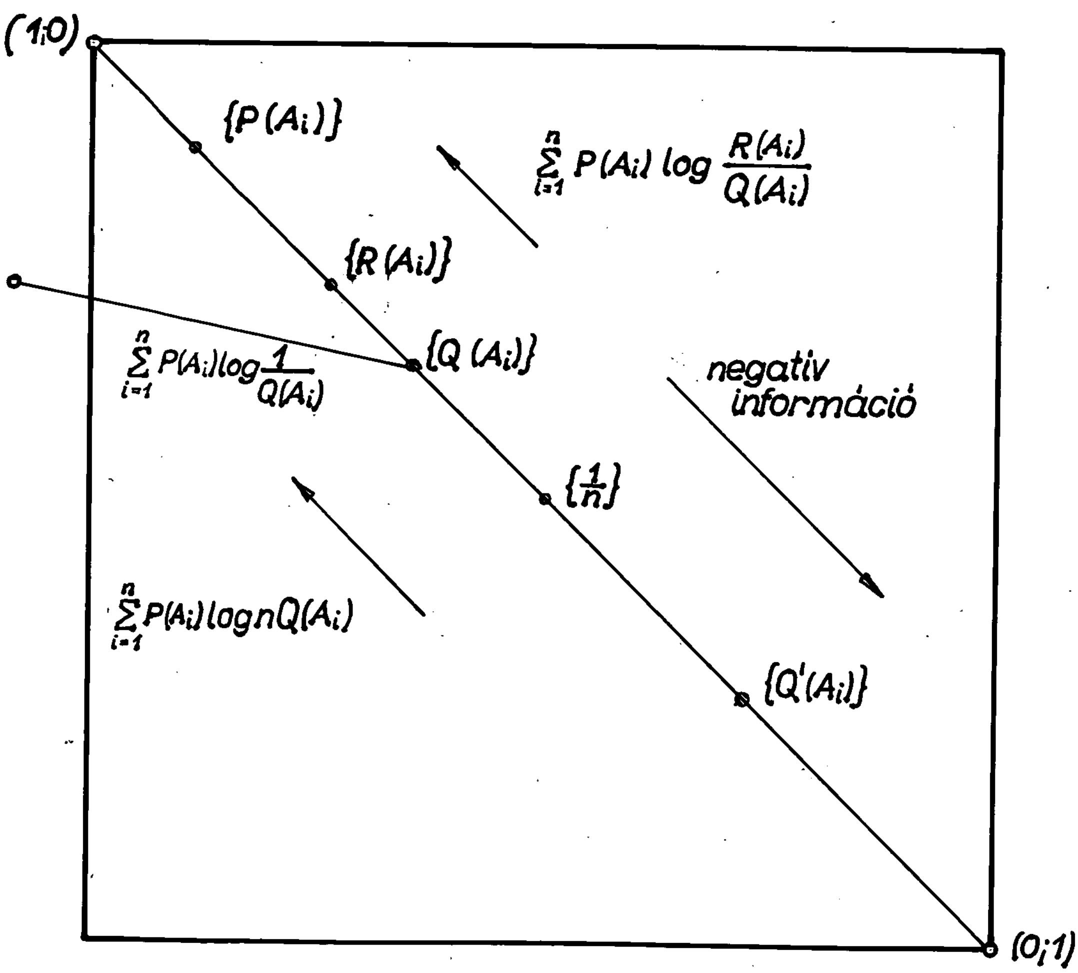
Ezt az eredményt hétköznapi nyelven ugy fogalmazhatjuk meg, hogy jobb, ha egy személlyel nem közlünk semmit, mint ha félrevezetjük. A félrevezetés ugyanis negativ értékü információként jelentkezik.

Másik érdekes tulajdonsága a definiciónak, hogy az egyenletes matematikai eloszlásu eseményrendszer esetén nincs szükség információra ahhoz, hogy a szubjektiv eloszlás megegyezzen a matematikaival, mert az ember eleve az egyenletes eloszlásból indul ki. Ha a matematikai formuláktól elszakadva vizsgáljuk ezt a kérdést, akkor is elfogadhatjuk ezt a következményt. Nincs semmi információ abban, ha valakivel azt közöljük, hogy egyformán gondolhat bármelyikre azok közül a válaszok közül, amelyeket lehetségesnek tart. Kivéve természetesen azt az esetet, amikor valamilyen negativ értékü információt kapott előtte a szubjektum. Ekkor pozitiv értékü információra van szükség ahhoz, hogy visszaálljon az eredeti egyenletes szubjektiv eloszlás.

A szubjektiv valószinüségeloszlás leirásánál /ld. [16.]/ kihangsulyoztuk a szubjektiv valószinüség statisztikai jellegét /magának a szubjektiv valószinüségnek is van szubjektiv eloszlása/. Ilyen szemszögből tekintve az általunk használt $Q(A_i)$ egy valószinüségi változó várható értéke, tehát $\log\frac{1}{Q(A_i)}$ ugyancsak várható értékként kezelendő. E várható értékekkel való számolás reális eredményt ad mindaddig, mig a $Q(A_i)$ -hez tartozó valószinüségi változó szórása kicsi.

Ellenkező esetben e statisztikai jellemzőket figyelembe kell venni. A szubjektiv valószinüség e tulajdonsága maga után vonja, hogy az információ értéke is rendelkezik e statisztikai tulajdonsággal, $\emptyset_A(I)$ szintén várható értékként kezelhető. E várható érték realitását ugyancsak az előbb emlitett valószinüségi változók szórása határozza meg. E kérdéskör részletes vizsgálata nem célunk, alapfeladatunk a megfelelő definiciók bevezetése, de a fentiek figyelembevételét fontosnak tartjuk az alkalmazás során.

Az elmondottak szemléletesebbé tételét kivánjuk elérni az alábbi 1. ábra segitségével.



1.ábra

A valószinüségeloszlásokat kétdimenzióban ábrázoltuk /ebben az esetben az $y = -x + 1$ egyenes első negyedbe eső szakaszán helyezkednek el az eloszlásokat ábrázoló pontok/, de a jelölésekkel az általános esetet akartuk érzékeltetni. Az üres karikák az elfajult eloszlásokat jelképezik /az 5 dimenziós térben például 5 üres karika lenne, ezek közül az egyik koordinátái $(0; 0; 0; 1; 0)$ /. A középpontban levő $\left\{\frac{1}{n}\right\}$ eloszlás a teljes bizonytalansági állapotot fejezi ki, a $\{Q\ (A_i)\}$, $\{R\ (A_i)\}$ eloszlások egyre közelebb vannak a $\{P\ (A_i)\}$ matematikai eloszláshoz, ami pozitiv információfelvételt jelent. A $\{Q'\ (A_i)\}$ eloszlás még távolabb van a $\{P\ (A_i)\}$-tól, mint az $\left\{\frac{1}{n}\right\}$, ezért ebbe az állapotba jutás negativ információfelvételt jelent. Az üres karika felé mutató nyil, illetve a ráirt mennyiség annak az információnak az értékét jelenti, amikor közöljük a szubjektummal, hogy melyik esemény következett be. Ez utóbbi magyarázatát később adjuk meg.

Fel kell még hivnunk a figyelmet arra, hogy két szubjektiv valószinüségeloszlás közül nem mindig az adja a kisebb szubjektiv entrópiát, amelyik közelebb áll a matematikai eloszláshoz. Ezt megint egy ellenpéldán mutatjuk be:

	A_1	A_2
$P\ (A_i)$	0,2	0,8
$Q_1(A_i)$	0,05	0,95
$Q_2(A_i)$	0,39	0,61

$\{Q_1(A_i)\}$ távolsága $\{P(A_i)\}$ -től:

$$T_1 = \sqrt{0{,}15^2 + 0{,}15^2} = 0{,}212 ,$$

$\{Q_2(A_i)\}$ távolsága $\{P(A_i)\}$ -től:

$$T_2 = \sqrt{0{,}19^2 + 0{,}19^2} = 0{,}262 .$$

$\{Q_1(A_i)\}$ által meghatározott szubjektiv entrópia:

$$0{,}2 \cdot \log \frac{1}{0{,}05} + 0{,}8 \cdot \log \frac{1}{0{,}95} = 0{,}923 .$$

$\{Q_2(A_i)\}$ által meghatározott szubjektiv entrópia:

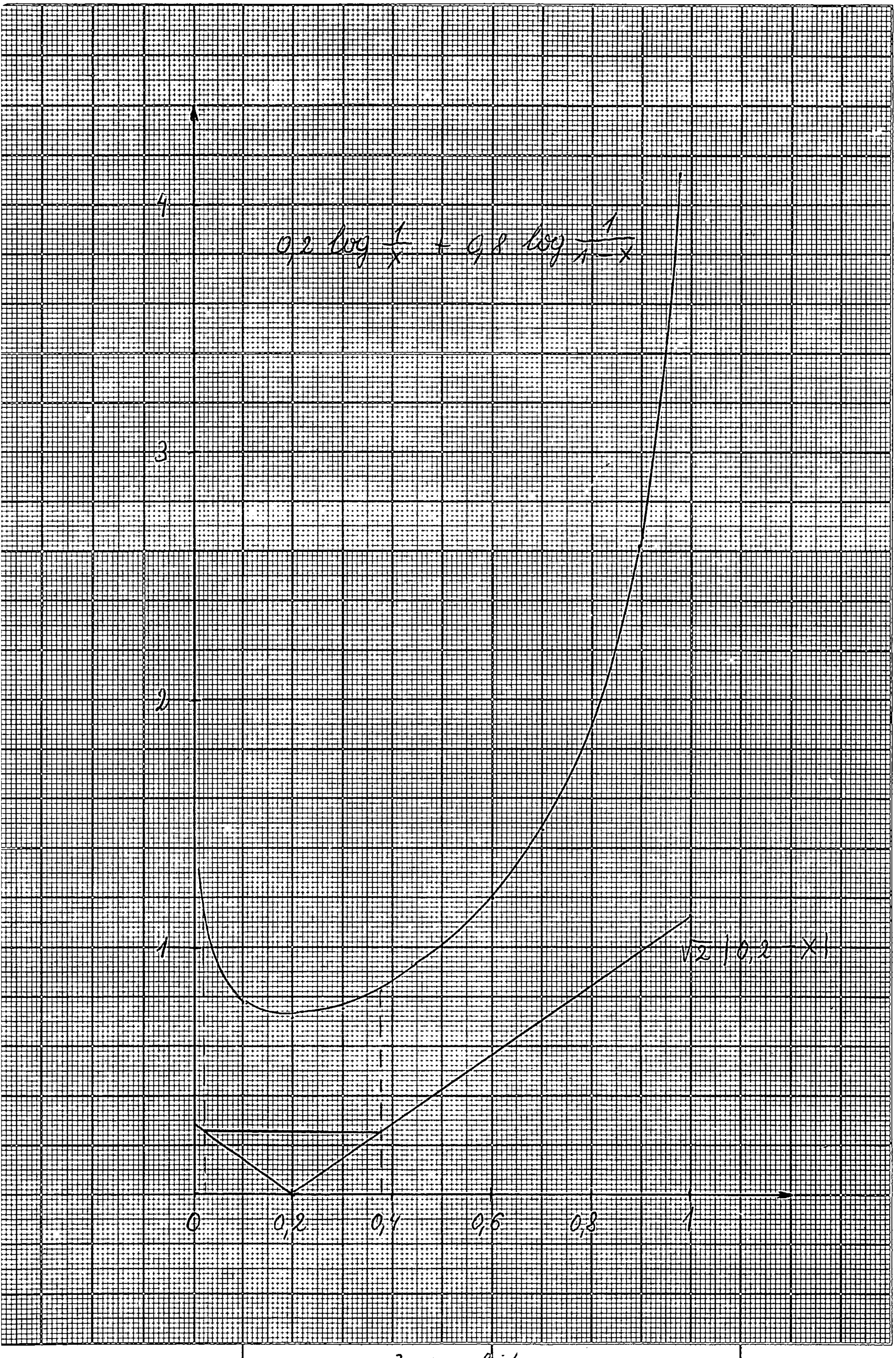
$$0{,}2 \cdot \log \frac{1}{0{,}39} + 0{,}8 \cdot \log \frac{1}{0{,}61} = 0{,}842 .$$

Tehát a $\{Q_1(A_i)\}$ szubjektiv eloszlás közelebb áll a matematikai eloszláshoz, mint a $\{Q_2(A_i)\}$ eloszlás, mégis az elsőnek nagyobb a szubjektiv entrópiája.

Ezt az érdekes, talán furcsa tulajdonságot ugyan ismernie kell e témával foglalkozó kutatónak, lényegének feltárása viszont tisztán matematikai probléma, ezzel itt nem foglalkozunk, csupán egy speciális esetben igyekszünk rámutatni ennek okára a 3. grafikon alapján.

Vegyünk egy eseményrendszert, mely két eseményből áll, és teljesülnek a következő feltételek:

$\mathcal{A}$	A_1	A_2
$P(A_i)$	0,2	0,8
$Q(A_i)$	x	1-x



3. grafikon

Ha felirjuk a szubjektiv entrópiát és a két eloszlás távolságát, akkor az alábbi két függvényhez jutunk:

$$x \longrightarrow 0{,}2 \log \frac{1}{x} + 0{,}8 \log \frac{1}{1-x} ,$$

$$x \rightarrow \sqrt{\left[0{,}2 - x\right]^2 + \left[0{,}8 - (1 - x)\right]^2} = \sqrt{2}\,\left|0{,}2 - x\right| .$$

E két függvényt ábrázoltuk a 3. grafikonon. Az x tengellyel párhuzamosan huzott zöld vonal két olyan pontot metsz ki a második függvény görbéjén, melyekhez tartozó szubjektiv eloszlások egyenlő távolságra vannak a matematikai eloszlástól. A hozzájuk tartozó, piros vonallal jelölt szubjektiv entrópiák viszont különbözőek. Az is látható, hogy a szóbanlevő két entrópia közti különbség annál nagyobb, minél közelebb van az egyik szubjektiv eloszlás a 0; 1 elfajult eloszláshoz. Képzeljük el ugyanis, hogy a zöld vonalat egy kicsit magasabban huzzuk meg, ez azt is jelenti, hogy x értéke közelebb került a 0-hoz, 1-x értéke pedig az 1-hez. Ebben az esetben a bal oldali piros vonal sokkal hosszabb lesz, a jobb oldali pedig alig változik, különbségük tehát rohamosan nő.

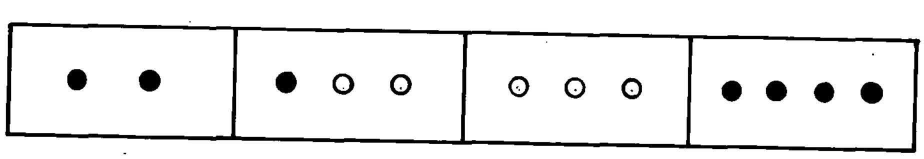
Ez a matematikai eredmény szintén helyesen tükröz pszichológiai tulajdonságot. Vegyük az ábrán kiemelt három eloszlást:

matematikai eloszlás:	0,2	0,8
szubjektiv eloszlás :	0,02	0,98
szubjektiv eloszlás :	0,38	0,62

A felvett két szubjektiv eloszlás egyenlő távol van ugyan a matematikaitól, informáltság tekintetében /a matematikai eredménytől eltekintve is/ mégis sokkal rosszabbnak kell venni az elsőt az miatt, hogy az szinte biztosra veszi a második esemény bekövetkeztét, ami persze nagyon rossz elképzelést takar, hisz a második esemény matematikai valószinüsége csak 0,8.

Az információ értékének általánosabb értelmezését néhány példa elemzésével kezdjük.

Vegyünk négy fiókot, ezekbe 12 számozott golyót helyezünk el, melyek közül 5 fehér és 7 fekete. Az első fiókba 2 fekete, a másodikba 1 fekete és 2 fehér, a harmadikba 3 fehér, a negyedikbe pedig 4 fekete golyót teszünk /a rájuk irt számok sorrendje lényegtelen az elosztásnál/.



2. ábra

Egy urnába helyezzünk el 1-től 12-ig számozott egyenlő nagyságu lapokat, huzás céljából. Az urnából kihuzunk egy számot. Első kérdés, hogy milyen szinü golyó számát huztuk ki.

$\mathcal{A}$: A_1 : fehér golyó számát huztuk A_2 : fekete golyó számát huztuk

$$P(A_1) = \frac{5}{12} \qquad P(A_2) = \frac{7}{12}$$

/A kedvező esetek számát osztjuk az összes lehetőségek számával./

$$Q(A_1) = 0{,}3 \qquad Q(A_2) = 0{,}7$$

/Ezek fiktiv számok./

Ha azt közöljük a v.sz.-lyel, hogy eddig 10-szer végeztük el a kisérletet ugy, hogy a kihuzott számot minden alkalommal visszatettük és mindig fehér golyót jelölő számot huztunk, e közlés valószinüleg megváltoztatja a szubjektiv eloszlást. Előfordulhat, hogy e közlés olyan érzést teremthet a v.sz.-ben, hogy a fehér golyónak nagyobb a valószinüsége, de az is előfordulhat, hogy ugy gondolkodik, hogy ennyi fehér után már "biztosan" fekete következik, tehát ennek ad nagyobb valószinüséget. Természetesen e közlés nem változtatja meg az eseményrendszert, sem annak matematikai eloszlását, igy az információ értékét a már leirt módon számithatnánk.

Más a helyzet, ha azt közöljük a v.sz.-lyel, hogy a kihuzott számu golyó nincs a 3. fiókban. E közléssel az eseményrendszer nem változott, mert továbbra is fehér és fekete golyók jöhetnek számitásba.

Tegyük fel, hogy a szubjektiv valószinüségfelosztás a következőre változott:

$$R(A_1) = 0{,}15 \qquad R(A_2) = 0{,}85$$

Könnyen belátható, hogy az uj szubjektiv entrópiát helytelen lenne az eredeti matematikai eloszlással számolni, vagyis az uj szubjektiv entrópia nem

$$\frac{5}{12} \cdot \log \frac{1}{0{,}15} + \frac{7}{12} \cdot \log \frac{1}{0{,}85}$$ lesz.

Ennek indoklására képzeljük el, hogy egy másik v.sz.-t hivunk a kisérlethez, akivel viszont csak az uj feltételeket ismertetjük. Mivel számára csak 9 golyó jöhet számitásba, ezért az ő esetében a

$$P'(A_1) = \frac{2}{9} \qquad P'(A_2) = \frac{7}{9}$$ matematikai eloszlással kell számolni. Ha most az ő szubjektiv eloszlása megegyezik az előző v.sz. uj /az információfelvétel utáni/ szubjektiv entrópiával, vagyis

$$Q'(A_1) = R(A_1) = 0{,}15 \quad \text{és} \quad Q'(A_2) = R(A_2) = 0{,}85 \;,$$

ekkor a második v.sz. szubjektiv entrópiája

$$\frac{2}{9} \cdot \log \frac{1}{0{,}15} + \frac{7}{12} \cdot \log \frac{1}{0{,}85} \;.$$

Ezek szerint viszont az eredeti vizsgálati személy esetében is ezt kell vennünk uj matematikai valószinüségeloszlásnak, hisz ellentmondáshoz jutnánk, ha két v.sz. ugyanazt az eseményrendszert figyelve ugyanazon szubjektiv eloszlással bir, és mi különböző szubjektiv entrópiaértékhez jutnánk a számolás során.

Az információ értékét tehát e példa esetében a következőképpen kell számolni:

$\mathscr{A}$	A_1 /fehér/	A_2 /fekete/
Eredeti matematikai eloszlás:	$\frac{5}{12}$	$\frac{7}{12}$
Eredeti szubjektiv eloszlás:	0,3	0,7

Eredeti szubjektiv entrópia:

$$\frac{5}{12} \cdot \log \frac{1}{0,3} + \frac{7}{12} \log \frac{1}{0,7} = 1,024 \,.$$

I : "Nem a harmadik fiókban levő golyó számát huztuk."

Uj matematikai eloszlás:	$\frac{2}{9}$	$\frac{7}{9}$
Uj szubjektiv eloszlás:	0,15	0,85

Uj szubjektiv entrópia:

$$\frac{2}{9} \cdot \log \frac{1}{0,15} + \frac{7}{9} \cdot \log \frac{1}{0,85} = 0,79 \,.$$

A felvett információ értéke a két szubjektiv entrópia különbsége:

$$\varnothing_{\mathscr{A}}(I) = 1,024 - 0,79 = 0,234 \,.$$

E példa alapján áttérhetünk az ilyen tipusu információfelvétel leirására tetszőleges $\mathscr{A}$ eseményrendszer esetében.

Legyen $\mathscr{A} = \{A_i\}$, ennek eredeti matematikai eloszlása $\{P\,(A_i)\}$, szubjektiv eloszlása pedig $\{Q\,(A_i)\}$. Tegyük fel, hogy I információ felvétele után uj $\{P'(A_i)\}$ matematikai és $\{Q'(A_i)\}$ szubjektiv eloszlásokkal kell számolnunk. Minden esetben i értéke 1-től n-ig mehet.

Az információ értékét most is a két szubjektiv entrópia különbségeként kapjuk:

$$\not{\mathcal{A}}(I) = \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} - \sum_{i=1}^{n} P'(A_i) \log \frac{1}{Q'(A_i)} .$$

Elvi szempontból fontos megjegyezni, hogy itt nem arról van szó, hogy a matematikai eloszlás megváltozik /ilyen a matematikában nem lehetséges/, hanem az információfelvétel után egy uj matematikai valószinüségi mezővel kell dolgoznunk. A már bevezetett formulával

$$P \times Q \times \mathcal{A} \xrightarrow{I} P' \times Q' \times \mathcal{A}$$

alakban fejezhetjük ki e folyamatot.

Hogy a legáltalánosabb esetet előkészitsük, térjünk vissza az előbbi kisérlethez. A v.sz.-nek most nem a golyó szinét kell figyelnie, hanem azt, hogy a kihuzott szám melyik fiókban levő golyót fogja jelölni.

A szóbajöhető események halmazát most $\mathcal{B}$ -vel jelöljük, ennek egy B_i eleme azt az eseményt jelöli, hogy a kihuzott szám az i-edik fiókban levő golyót határozza meg.

Matematikai valószinüségek:

$$P(B_1) = \frac{2}{12} \; ; \quad P(B_2) = \frac{3}{12} \; ; \quad P(B_3) = \frac{3}{12} \; ; \quad P(B_4) = \frac{4}{12}$$

Szubjektiv valószinüségek /fiktiv adatok/:

$$Q(B_1) = 0{,}15 \; ; \quad Q(B_2) = 0{,}2 \; ; \quad Q(B_3) = 0{,}2 \; ; \quad Q(B_4) = 0{,}45$$

Szubjektiv entrópia:

$$\frac{2}{12}\log\frac{1}{0{,}15}+\frac{3}{12}\log\frac{1}{0{,}2}+\frac{3}{12}\log\frac{1}{0{,}2}+\frac{4}{12}\log\frac{1}{0{,}45}=2.$$

Tegyük fel, hogy most is ugyanazt az I információt közöljük: "a kihuzott számu golyó nincs a harmadik fiókban".

Ebben az esetben már a lehetséges esem[illegible]yek halmaza is megváltozott, hisz B_3 lehetőség kiesett.

Uj eseményrendszer $/\mathcal{B}'/$: B_1' ; B_2' ; B_3'

Az uj matematikai eloszlás:

$$P'(B_1') = \frac{2}{9} \;;\quad P'(B_2') = \frac{3}{9} \;;\quad P'(B_3') = \frac{4}{9} \;.$$

Uj szubjektiv eloszlás /fiktiv adatok/:

$$Q'(B_1') = 0{,}2 \;;\quad Q'(B_2') = 0{,}29 \;;\quad Q'(B_3') = 0{,}51 \;.$$

Uj szubjektiv entrópia:

$$\frac{2}{9}\log\frac{1}{0{,}2}+\frac{3}{9}\log\frac{1}{0{,}29}+\frac{4}{9}\log\frac{1}{0{,}51} = 1{,}542 \;.$$

A felvett információ értéke:

$2 - 1{,}542 = 0{,}458$, most is a két szubjektiv entrópia különbsége.

E két számpélda összehasonlitásakor szépen látható az a már leirt véleményünk helyessége, hogy egy adott közlés két eseményrendszer esetén különböző információértéket adhat. I mindkét esetben ugyanazon közlés eredménye, mégis

$\emptyset_{\mathcal{A}}(I) = 0{,}234$ és $\emptyset_{\mathcal{B}}(I) = 0{,}458$ -nek adódott.

Ebből az is látható, hogy a közlés /általában az észlelés/ fogalma lényegesen különbözik az információ fogalmától.

Ezzel elérkeztünk a legáltalánosabb esethez, amikor egy adott információ megváltoztatja a figyelembeveendő eseményrendszert, vele együtt a matematikai és szubjektiv valószinüségi mezőt.

Algebrailag mindez a következőképpen irható le:

Eredeti eseményrendszer $/\mathcal{A}/$: $\{A_i\}$

Eredeti matematikai valószinüségeloszlás : $\{P(A_i)\}$

Eredeti szubjektiv valószinüségeloszlás : $\{Q(A_i)\}$

Eredeti szubjektiv entrópia: $\sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)}$

Minden esetben i = 1, 2, ..., n.

I információ felvétele után:

Uj eseményrendszer $/\mathcal{A}'/$: $\{A_j'\}$

Uj matematikai valószinüségeloszlás: $\{P'(A_j')\}$

Uj szubjektiv valószinüségeloszlás: $\{Q'(A_j')\}$

Uj szubjektiv entrópia: $\sum_{i=1}^{m} P'(A_j') \log \frac{1}{Q'(A_j)}$

Minden esetben i = 1, 2, ..., m.

A felvett információ értéke:

$$\emptyset_{\mathcal{A}}(I) = \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} - \sum_{j=1}^{m} P'(A_j) \log \frac{1}{Q'(A_j)}$$

A folyamat formulája:

$$P \times Q \times \mathcal{A} \xrightarrow{I} P' \times Q' \times \mathcal{A}'$$

Ez az általános definició természetesen alkalmazható speciális esetekre is /például amikor a matematikai eloszlás nem változik/ és ugyanazt az eredményt adja, mint az előző definiciók.

Ha a már emlitett $0 \cdot \log \frac{1}{0} = 0$ definiciót elfogadjuk, akkor $\mathcal{A}$ eseményrendszer megváltoztatásának nincs jelentősége. Ekkor ugyanis az előbbi esetet ugy is felfoghatjuk, hogy $\mathcal{A}$ némely eseményének matematikai és szubjektiv valószinüsége egyszerre lesz 0, ami maga után vonja, hogy a szubjektiv entrópia kiszámitásánál a megfelelő tagok is nullák lesznek, igy a felfogástól függetlenül ugyanarra az eredményre jutunk. Mindez indokolja különben e fejezet cimét is, hogy rögzitett $\mathcal{A}$ -ra vonatkozó információfelvétellel dolgozunk, és a $\emptyset_{\mathcal{A}}(I)$ jelölésnél sem kell figyelembevennünk $\mathcal{A}'$ -t.

Most már indokolhatjuk az 1. ábrán szereplő

$$\sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)}$$ információmennyiséget, melyet

az üres karika felé mutató nyilra irtunk. Ha a szubjektummal közöljük a végrehajtott kisérlet eredményét, hogy konkrétan melyik esemény következett be, akkor az uj eseményrendszer egy eseményből fog állni, melynek szubjektiv és matematikai valószinüsége 1. Igy a fenti képletbe helyettesitve a

$$\emptyset_{\mathcal{A}}(I) = \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} - 1 \cdot \log \frac{1}{1} =$$

$$= \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)}$$ eredményre jutunk.

Tehát a konkrét kimenetelt közölve az információ értéke éppen a kezdeti szubjektiv entrópiával egyezik meg. /Vigyázat ezzel nem tagadtuk meg e fejezet elején tett kijelentésünket, miszerint A_i esemény bekövetkeztének közlésekor $\log \frac{1}{Q(A_i)}$ meglepetést okozunk a szubjektumnak !/

III. Az információ értéke adott szubjektumra és rögzitett ξ valószinüségi változóra vonatkozóan.

Mint a cimből is kitünik, e fejezet tartalmilag nem fog különbözni az előbbitől, az eddig bevezetett fogalmakat fogjuk átvinni valószinüségi változóra.

Feladatunk annyiban lesz könnyebb, hogy nem kell uj fogalmakat leirnunk, viszont mélyebb matematikai apparátusra kell támaszkodnunk, aminek magyarázata kissé nehezebb lesz, kevésbé tudjuk tartani magunkat ahhoz az álláspontunkhoz,hogy matematikailag egzaktan irjuk le mondanivalónkat, ugyanakkor közérthető is legyen. Esetenként valamelyik feltételből engednünk kell.

Kezdjük a valószinüségi változó fogalmával.

Legegyszerübben ugy mondhatnánk, hogy ha a lehetséges események számok, akkor ezeket a számokat valószinüségi változóknak nevezzük.

Nézzünk egy-két példát: Gondoljunk vissza a "fiókos" kisérletre /2. ábra/. Tegyük fel, hogy az egyes fiókokat sorszámoztuk, ekkor kérdezhetjük, hogy az urnából kihuzott szám hanyadik fiókban levő golyót adja. Négy lehetőségünk van:

$x_1 = 1$; $x_2 = 2$; $x_3 = 3$; $x_4 = 4$. Itt a ξ valószinüségi változó értékei 1; 2; 3; 4 lehetnek.

Ha két kockával dobunk egyszerre, és a felülre került számok összegét figyeljük, akkor a valószinüségi változó értéke 2; 3; ...; vagy 12 lehet.

E példák /mivel az előző fejezetben is szerepeltek/ csak a szemléletváltozást kivánjuk bemutatni. A szubjektiv entrópia kiszámitásánál mindegy, hogy azt mondjuk, hogy A_3 esemény következett be, vagy ha $\xi = 3$-at közöljük, mivel a valószinüségértékek mindkét esetben ugyanazok.

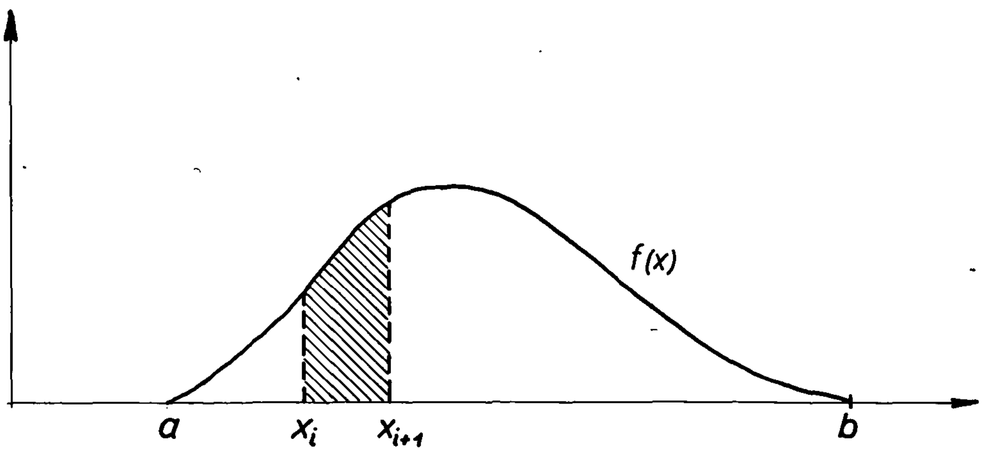
Lényegesen más a helyzet, ha azt mondjuk a v.sz.-nek, hogy "az utcán találtam egy ceruzát, mondja meg milyen hosszu". Egy ceruza hosszuságának értékére végtelen sok lehetőség létezik /tegyük fel, hogy 0 és 20 cm között változhat/. Véletlen, hogy milyen hosszu ceruzát találunk, tehát a talált ceruza hosszuságának lehetséges értékeit valószinüségi változónak nevezzük, és általában ξ vagy η görög betükkel jelöljük. Az utóbbi példa esetében ξ értéke 0 és 20 között változhat.

Az első két példában bemutatott valószinüségi változókat diszkrét valószinüségi változóknak, az utóbbit pedig folytonos valószinüségi változónak nevezzük. A diszkrét valószinüségi változóra lehetne még az előbbitől olyan eltérő példát is hozni, amikor az is végtelen sok értéket vehet fel, erre azonban itt nem térünk ki, hisz a dolgozat pszichológiai oldalát nem érinti, esetleges gyakorlati előfordulásakor csupán matematikai problémák léphetnek fel azon tul, amit itt megtárgyalunk.

Mivel a diszkrét valószinüségi változó esetében a második fejezetben leirt módon lehet meghatározni a szubjektiv entrópiát, és ez alapján az információfelvétel értékét, ezért ezzel már nem foglalkozunk.

A folytonos valószinüségi változó esetében általában egy-egy konkrét érték valószinüsége 0. Például 0 annak a matematikai valószinüsége, hogy éppen 12,25 cm hosszu ceruzát találunk vagy, hogy a céltábla középétől 4,1 cm-re levő pontba esik a lövés. /Mégegyszer hangsulyozzuk, hogy a 0 valószinüségü esemény nem mindig lehetetlen./ Az információfelvételt tehát nem mérhetjük a már megismert képlettel folytonos valószinüségi változó esetében. Ehhez először a matematikai és a szubjektiv sürüségfüggvényt kell leirni.

Vegyünk fel egy derékszögü koordinátarendszert, ennek x tengelyén jelöljük ξ lehetséges értékeit.
Tegyük fel, hogy ezek az értékek a-tól b-ig mehetnek.
/3.ábra./



3.ábra

Osszuk ezt az $[a;\ b]$ intervallumot n egyenlő részre, az osztáspontok $a = x_o$; x_1; ...; $x_n = b$ legyenek. Rajzoljunk az $[a;\ b]$ intervallum fölé olyan /függvényt ábrázoló/ vonalat, hogy annak x_i-től x_{i+1} -ig terjedő része alatti terület annak valószinüségével egyezzék meg, hogy az adott valószinüségi változó az $[x_i;\ x_{i+1}]$ intervallumba esik, bárhogyan vesszük fel x_i -t és x_{i+1} -et. A 3. ábrán besatirozott rész jelöli e területet. Mivel ξ értékei biztosan a és b közé esnek, ezért a görbe alatti területnek a-tól b-ig 1-nek kell lenni. E görbe által meghatározott függvényt a valószinüségi változó sürüségfüggvényének nevezzük.

Mivel a görbe alatti területet integrálszámitással határozzuk meg, ezért annak valószinüsége, hogy a ξ x_i és x_{i+1} közé esik

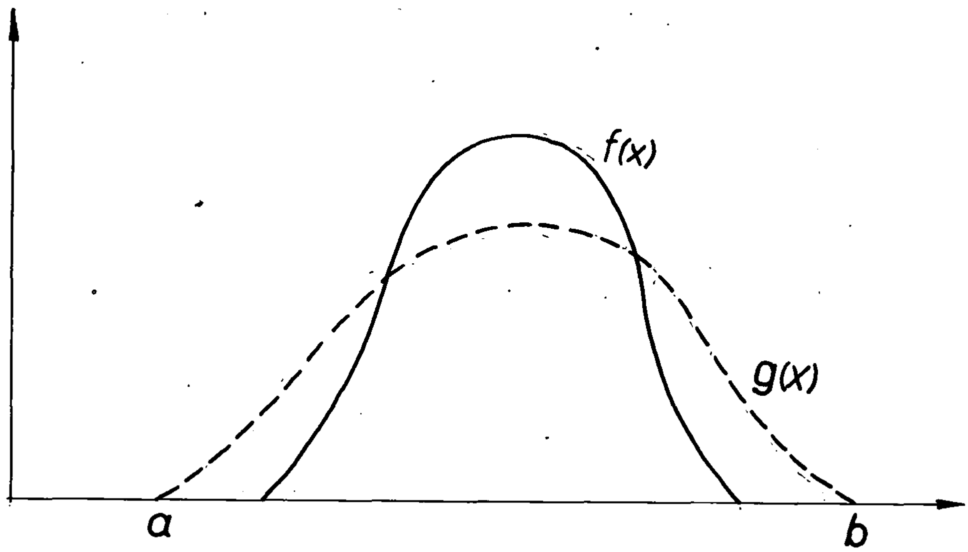
$$\int_{x_i}^{x_{i+1}} f(x)\, dx \quad , \quad \text{jelekkel}$$

$$P\left(x_i \leq \xi \leq x_{i+1}\right) = \int_{x_i}^{x_{i+1}} f\ x\ dx\ .$$

$g\ (x)$-el jelöljük a szubjektiv sürüségfüggvényt, amely ugyanilyen módon határozza meg annak szubjektiv valószinüségét, hogy ξ x_i és x_{i+1} közé esik, tehát

$$Q\left(x_i \leq \xi \leq x_{i+1}\right) = \int_{x_i}^{x_{i+1}} g\ (x)\ d\ x\ .$$

A g(x) értelmezési tartománya /az az intervallum, ahol g (x) felveszi értékeit/ magába kell, hogy foglalja f (x) értelmezési tartományát /ld. 4.ábra/.



4.ábra

Ha e kikötést nem tennénk, akkor előállhatna az a helyzet, hogy egy részintervallum szubjektiv valószinüsége 0, matematikai valószinüsége pedig attól különböző, igy /mint már a második fejezetben indokoltuk/ a szubjektiv entrópiát nem tudnánk mérni.

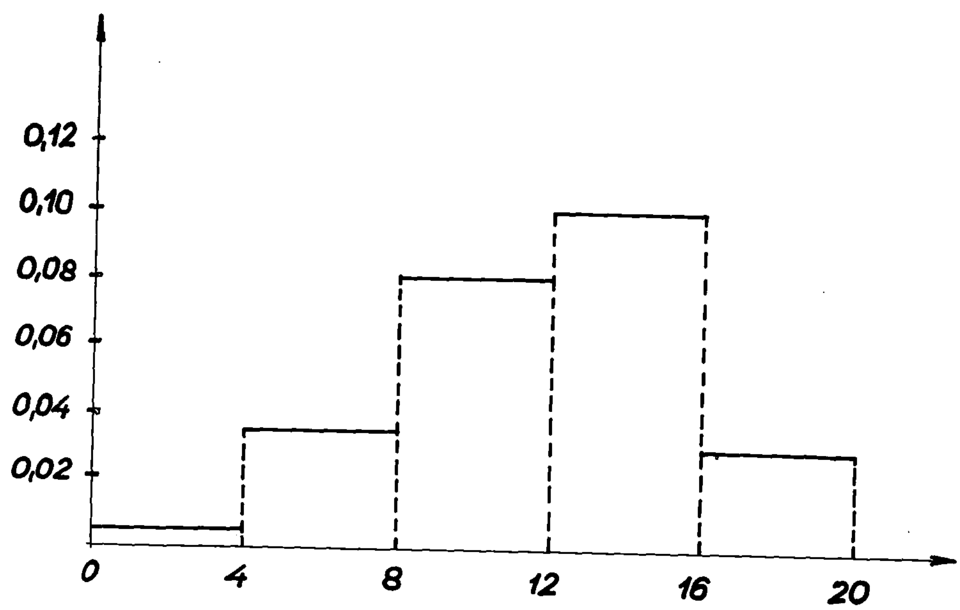
A matematikai sürüségfüggvény vagy adott, vagy a feltételekből, esetleg mérés utján kell meghatározni. Itt a harmadik eset tárgyalásának van értelme.

Mérés utján az ugynevezett emprikus matematikai sürüségfüggvényt tudjuk meghatározni, ami általában csak közeliti a valódi sürüségfüggvényt. A közelités pontossága azon mulik, milyen részletes méréseket végzünk.

Ha például egy talált ceruza hosszának, mint valószinüségi változónak a matematikai sürüségfüggvényéről akarunk képet kapni, a következőképpen járhatunk el: Megmérjük egy iskolai tanuló által használt ceruzák hosszát. Tegyük fel, hogy 300 tanuló esetében az alábbi eredményt kapjuk:

hosszuság /cm/	0-4	4-8	8-12	12-16	16-2o
ceruzák száma	5	42	94	122	37

Ekkor 1 ceruzára $\frac{1}{4 \cdot 300}$ magasságot kell számitanunk /mivel minden téglalap területét ugy kell kiszámitani, hogy 4el kell szorozni a magasságot/, igy a téglalapok területének összege 1 lesz /5.ábra/.



5. ábra

Ha kisebb intervallumokra végezzük el a mérést, akkor pontosabban közelitjük az elméleti sürüségfüggvényt.

Itt szépen kirajzolódik az a probléma, amit a matematikai valószinüség objektivitásáról mondtunk.

Vajon a fenti grafikon valóban egy elvesztett ceruza hosszának, mint valószinüségi változónak a sürüségfüggvényét közeliti? Ez csak a következő /eddig elhallgatott/ feltételek mellett igaz:

1. Csak a tanulók által elvesztett ceruzákat vesszük figyelembe.
2. Ebben az iskolában felmért adatok jellemzik az adott város összes tanulóit.

Ha e feltételeket nem fogadjuk el, akkor a méréseket is általánosabb feltételek mellett kell elvégeznünk.

Esetenként szinte lehetetlen meghatározni a matematikai sürüségfüggvényt. Ha például a tanár megkérdezi a tanulótól, hogy normális körülmények között mennyi ideig él egy galamb, akkor a galamb életkorát valószinüségi változónak kell venni, hisz a tanár sem tud erre pontos választ adni. Sőt valószinü, hogy a matematikai sürüségfüggvényt sem ismeri, talán fel sem mérték elegendő pontossággal. Hogyan lehet ekkor eldönteni a válasz helyességét? Arra természetesen nincs módja a tanárnak, hogy a szükséges /objektive létező/ sürüségfüggvényt meghatározza. Nincs más hátra, mint a tanárban fellépő /különben szubjektiv/ eloszlást vegyük matematikainak. Ennek jogosságát a tanár tapasztalata, felkészültsége indokolja.

A felméréssel kapcsolatban meg kell még jegyezni, hogy folytonos valószinüségi változó esetében /vagy ha a lehetséges események száma nagy/ nem lehet az összes lehetséges esetet felmérni, hanem ugynevezett mintavételt végzünk, és ez alapján határozzuk meg a sürüségfüggvényt. E kérdéssel a matematikai statisztika foglalkozik.

A szubjektiv sürüségfüggvényt az előbbihez hasonlóan /természetesen most már csak az adott szubjektum véleményét véve figyelembe/ határozhatjuk meg. Erre konkrét, valódi adatokkal kidolgozott példát a már többször emlitett [16] dolgozatunkban találhatunk.

A továbbiakban feltételezzük a matematikai és a szubjektiv sürüségfüggvények folytonosságát. Ez a feltétel nem játszik különösebb szerepet, de igy az integrálhatóságot biztositottuk.

Térjünk rá ezek után a szubjektiv entrópia, majd a felvett információ mérésére folytonos valószinüségi változó esetében.

Legyen ξ egy folytonos valószinüségi változó, ennek $[a; b]$ -on értelmezett $f(x)$ matematikai sürüségfüggvénye és az $[a; b]$ -ot tartalmazó intervallumon értelmezett $g(x)$ szubjektiv sürüségfüggvénye.

Osszuk az $[a; b]$ intervallumot n egyenlő részre. Felfoghatjuk ugy is, hogy ezzel n eseményből álló eseményrendszert definiáltunk: A_i esemény akkor következik be, ha ξ az $[x_i; x_{i+1}]$ intervallumba esik.

Mivel

$$Q\left(x_i \leq \xi \leq x_{i+1}\right) = \int_{x_i}^{x_{i+1}} g(x)\,dx \quad \text{, ezért}$$

ha a szubjektummal azt közöljük, hogy ξ értéke x_i és x_{i+1} közé esett, akkor

$$\log \frac{1}{\int_{x_i}^{x_{i+1}} g(x)\,dx} \qquad \text{meglepetést okozunk.}$$

Ennek matematikai valószinüsége viszont

$$\int_{x_i}^{x_{i+1}} f(x)\, d\, x \;,\; \text{ezért a várható}$$

meglepetés, illetve az adott beosztáshoz tartozó szubjektiv entrópia

$$\sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)\, d\, x \cdot \log \frac{1}{\int_{x_i}^{x_{i+1}} g(x)\, d\, x} \;.$$

Lényeges, hogy ez az érték függ az $[a;\ b]$ intervallum beosztásától. Ez pszichológiai szempontból is érthető, mert annál nagyobb a szubjektum bizonytalansága, minél kisebb /ami maga után vonja, hogy több/ intervallumot jelölünk ki számára, amiket figyelembe kell vennie. Ez azt jelenti, hogy folytonos valószinüségi változóhoz az eddigi értelemben nem rendelhetünk egyértelmüen szubjektiv entrópiát, ez függ az intervallum felosztásától is. Könnyen belátható, hogy a beosztás finomitásával akármilyen nagy szubjektiv entrópia is felléphet. Pontosabban $n \longrightarrow \infty$ esetén a szubjektiv entrópia értéke is tart a végtelenhez.

Ennek ellenére folytonos valószinüségi változóhoz is definiálhatunk /a beosztástól független/ szubjektiv entrópiát, ugyanis az információelméletben egy $f(x)$ sürüségfüggvényü eloszlás entrópiáján az

$$\int_a^b f(x) \log \frac{1}{f(x)} d\, x \quad \text{-et értik.}$$

Ennek analógiájára a szubjektiv entrópiát az

$$\int_a^b f(x) \log \frac{1}{g(x)} \, d\,x$$ integrállal definiálhatnánk.

Ez bizonyos szempontból nem illeszkedik az eddigi definicióhoz, más jellegü mennyiség. Ezért számitásainknál a diszkrét esetben leirt definiciót használjuk, de az információ értékére ugyanazt az eredményt kapnánk akkor is, ha ezt a definiciót fogadnánk el a szubjektiv entrópia kiszámitására. Erre az adott esetben rá fogunk mutatni.

Az előző fejezetben ugy számitottuk ki a szubjektiv eloszlásban tartalmazott információ értékét, hogy az egyenlő szubjektiv valószinüségekkel számolt entrópiából kivontuk a tényleges szubjektiv valószinüségértékekkel számolt entrópiát /.40.oldal/. Ezzel analóg módon az egyenletes szubjektiv valószinüségeloszláshoz tartozó entrópiából kivonva az adott felosztáshoz tartozó tényleges szubjektiv entrópiát, megkapjuk az eloszlásban tartalmazott információ értékét az adott felosztásra.

Az egyenletes eloszlás itt azt jelenti, hogy egyenlő intervallumokba egyenlő valószinüséggel eshet ξ értéke. Az egyenletes eloszlás sürüségfüggvénye az x tengellyel párhuzamos, attól $\frac{1}{b-a}$ távolságra levő szakasszal ábrázolható, vagyis az $\frac{1}{b-a}$ konstans függvény.

Mivel
$$\int_{x_i}^{x_{i+1}} \frac{1}{b-a} \, d\,x = \frac{x_{i+1} - x_i}{b - a} ,$$

ezért az egyenletes eloszláshoz tartozó szubjektiv entrópia

$$\sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x) \, . \, \log \frac{1}{\frac{x_{i+1} - x_i}{b - a}} =$$

$$= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x) \log \frac{b - a}{x_{i+1} - x_i}$$

A két entrópia különbsége:

$$\sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x) \, d\, x \, . \, \log \frac{b - a}{x_{i+1} - x_i} -$$

$$\sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x) \, d\, x \, . \, \log \frac{1}{\int_{x_i}^{x_{i+1}} g(x) \, d\, x} =$$

$$= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x) d\, x \, . \, \log \frac{b - a}{x_{i+1} - x_i} \int_{x_i}^{x_{i+1}} g(x) d\, x \quad .$$

Az

$$\int_{x_i}^{x_{i+1}} f(x) d\, x = \left(x_{i+1} - x_i\right) f\left(z_i\right), \text{ ahol } x_i \leq z_i \leq x_{i+1}$$

és

$$\int_{x_i}^{x_{i+1}} g(x) \, d\, x = \left(x_{i+1} - x_i\right) g\left(z_i'\right), \text{ ahol } x_i \leq z_i' \leq x_{i+1}$$

egyenlőségeket felhasználva a fenti különbség

$$\sum_{i=0}^{n-1} \left(x_{i+1} - x_i\right) f\left(z_i\right) \, . \, \log \left(b - a\right) g\left(z_i'\right)$$ alakban irható,

aminek határértéke

$$\int_a^b f(x) \cdot \log (b - a) g(x) \, d x \quad , \text{ha} \quad n \to \infty.$$

E zárt alakban kapott kifejezés a $g(x)$, mint szubjektiv sürüségfüggvény által meghatározott, a szubjektiv eloszlásban tartámazott információ értékét jelenti, ami már az $[a; b]$ felosztásától független.

Vizsgáljuk meg ezek után azt az esetet, amikor egy információ hatására a $g(x)$, mint a priori szubjektiv sürüségfüggvény $r(x)$ pospriori szubjektiv sürüségfüggvénybe megy át. Az információ értékét az adott valószinüségi mezőre az előbbihez hasonló határátmenettel kapjuk:

$$\phi_{\xi}(I) = \int_a^b f(x) \cdot \log \frac{r(x)}{g(x)} \, d x .$$

Tehát a folytonos eloszlásu valószinüségi változó esetében is egy konkrét számot rendeltünk az információhoz, annak értékéül.

Ha a szubjektiv entrópiát az előbb emlitett

$$\int_a^b f(x) \cdot \log \frac{1}{g(x)} \, d x \quad \text{integrállal definiáljuk,}$$

akkor $\phi_{\xi}(I)$ előbbi értéke most is a két szubjektiv entrópia különbségeként adódik:

$$\phi_{\xi}(I) = \int_a^b f(x) \log \frac{1}{g(x)} \, d x - \int_a^b f(x) \log \frac{1}{r(x)} \, d x =$$

$$= \int_a^b f(x) \log \frac{r(x)}{g(x)} \, d x .$$

Meg kell jegyeznünk, hogy az információ értékére adott fenti zárt alak inkább csak matematikai szempontból érdekes. Amikor gyakorlati alkalmazásra kerül sor, általában az ott szereplő függvények /különösen a $g(x)$ és az $r(x)$/ nincsenek megadva képlettel, empirikus alakjuk ismert, tehát az integrál kiszámitása is közelitő módszerrel történik. A gyakorlati számolásnál az látszik legcélszerűbbnek, hogy a két szubjektiv entrópiát egy bizonyos beosztásra /mely beosztás már a $g(x)$ és az $r(x)$ közelitő meghatározásakor is szerepel/ számitjuk ki, igy a kettő különbsége adja $\phi_{\xi}(I)$ közelitő értékét:

$$\phi_{\xi}(I) \approx \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)\,dx \cdot \log \frac{1}{\int_{x_i}^{x_{i+1}} g(x)\,dx} -$$

$$- \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)\,dx \cdot \log \frac{1}{\int_{x_i}^{x_{i+1}} r(x)\,dx} =$$

$$= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)\,dx \cdot \log \frac{\int_{x_i}^{x_{i+1}} r(x)\,dx}{\int_{x_i}^{x_{i+1}} g(x)\,dx}$$

/Az itt szereplő integrálok kiszámitásával nincs dolgunk, mert a felmérés éppen ezek értékét adja./

Azt mondtuk, hogy ha a beosztást finomitjuk, akkor a szubjektiv entrópiák értékei határtalanul növekednek.

Ez nem okoz problémát, mert a különbségükre ez nem áll, azok egy adott szám körül ingadoznak, ahhoz egyre-egyre közelitenek, igy annál pontosabb eredményt kapunk $\phi_\xi(I)$ -re, minél nagyobbra választjuk n értékét.

Nézzünk egy példát: Vegyünk egy céltáblát, amelyen /egyenlőre/ nincsenek körgyürük. Lőjünk erre véletlenszerüen /ld: 18. oldal/. Jelentse ξ azt a számot, amilyen távolságra esik a lövés a kör középpontjától. Ha a kör sugara r, akkor az alábbi 4. grafikon f(x) közelitését adja, ha most is elfogadjuk ezt a feltételt, hogy egy lövés adott területbe való esésének valószinüsége a terület nagyságával egyenesen arányos. Az $[a; b]$ szerepét most a $[0; r]$ intervallum veszi át.

Osszuk fel ezt az intervallumot 10 egyenlő részre. A céltáblán ez a felosztás 10 körgyürünek /a legbelső kör/ felel meg.

Hogyan kell f(x) közelitő függvényét meghatározni?

Tegyük fel példaként a következő kérdést:

Mi a matematikai valószinüsége, hogy a 7. körgyürübe esik a lövés?

A 7. körgyürü belső sugara $\frac{6}{10}$ r, külső sugara $\frac{7}{10}$ r, területe

$$\left(\frac{7}{10} r\right)^2 \pi - \left(\frac{6}{10} r\right)^2 \pi = \frac{13}{100} r^2 \cdot \pi ,$$

tehát a teljes kör 0,13 része. Következik, 0,13 annak a valószinüsége, hogy a 7. körgyürübe esik a véletlen lövés.

f(x) -ről azt mondtuk, hogy olyan függvény, hogy a görbéje alatti területből a valószinüségértékekre lehet következtetni /ld. 61 oldal/.

Ahhoz, hogy az empirikus /lépcsős/ függvény alatti terület a 0,1 r hosszuságu intervallumon /ez az érték onnan jön, hogy az r hosszuságu intervallumot 10 egyenlő részre osztottuk/ a függvény alatti terület a már kiszámitott 0,13 legyen, a függvény értékét $\frac{1,3}{r}$-nek kell venni. Tehát a hetedik intervallumon a függvény értéke $\frac{1,3}{r}$. Hasonló számitással kapjuk az f(x) -et közelitő függvény értékeit, amiket az alábbi táblázatból olvashatunk ki.

inter-vallum	0 - 0,1r	0,1r-0,2r	0,2r-0,3r	0,3r-0,4r	0,4r-0,5r	0,5r-0,6r	0,6r-0,7r	0,7 r-0,8r	0,8r-0,9r	0,9r-r
f x	$\frac{0,1}{r}$	$\frac{0,3}{r}$	$\frac{0,5}{r}$	$\frac{0,7}{r}$	$\frac{0,9}{r}$	$\frac{1,1}{r}$	$\frac{1,3}{r}$	$\frac{1,5}{r}$	$\frac{1,7}{r}$	$\frac{1,9}{r}$

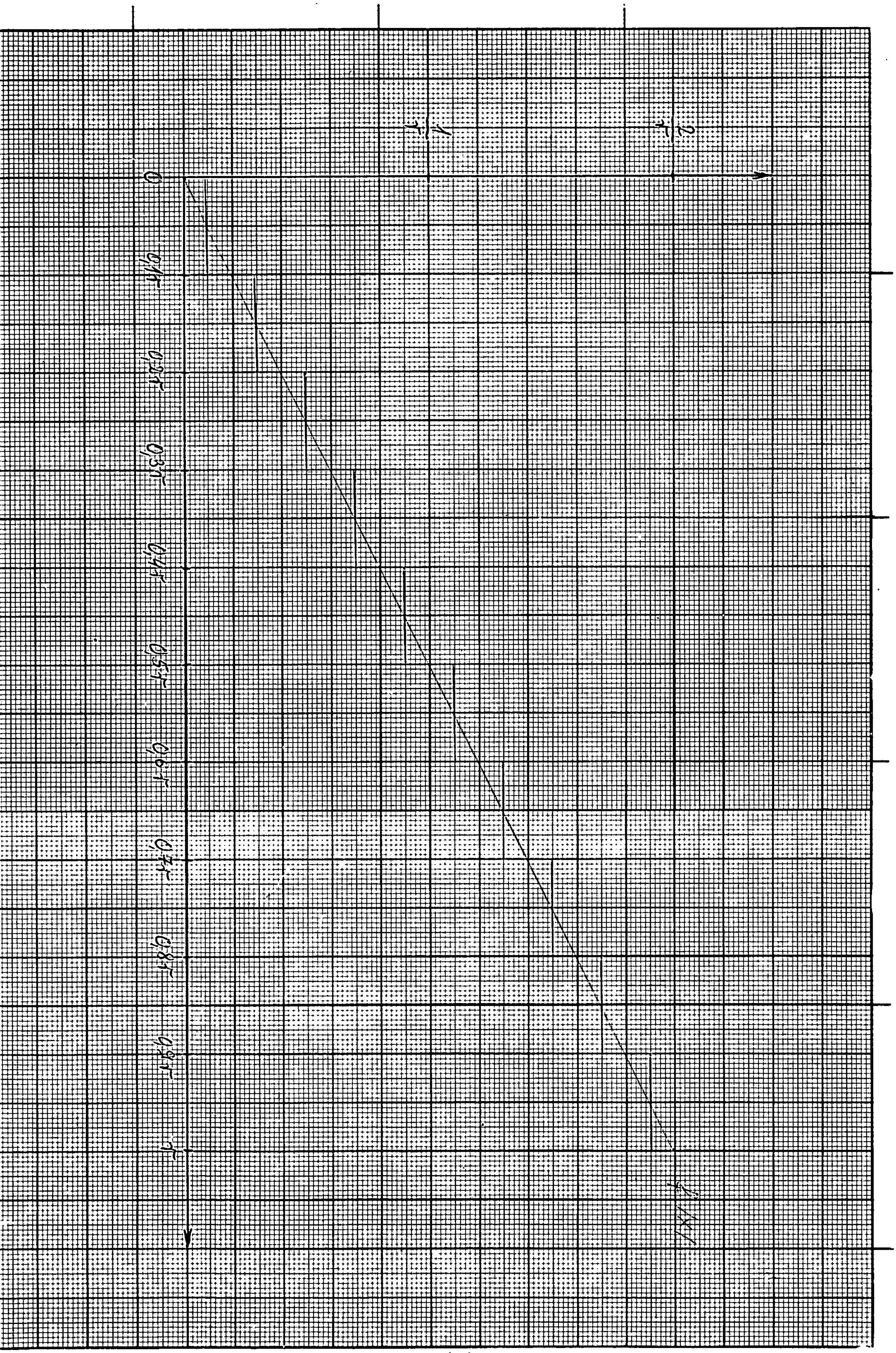
E táblázat alapján készitettük a 4. grafikont. Az f(x) közelitő függvényét $\bar{f}(x)$-el jelöltük.

Belátható, ha a beosztást finomitjuk, akkor ez a függvény egyre jobban közeliti a $\frac{2}{r^2} \cdot x$ függvényt, vagyis $f(x) = \frac{2}{r^2} x$.

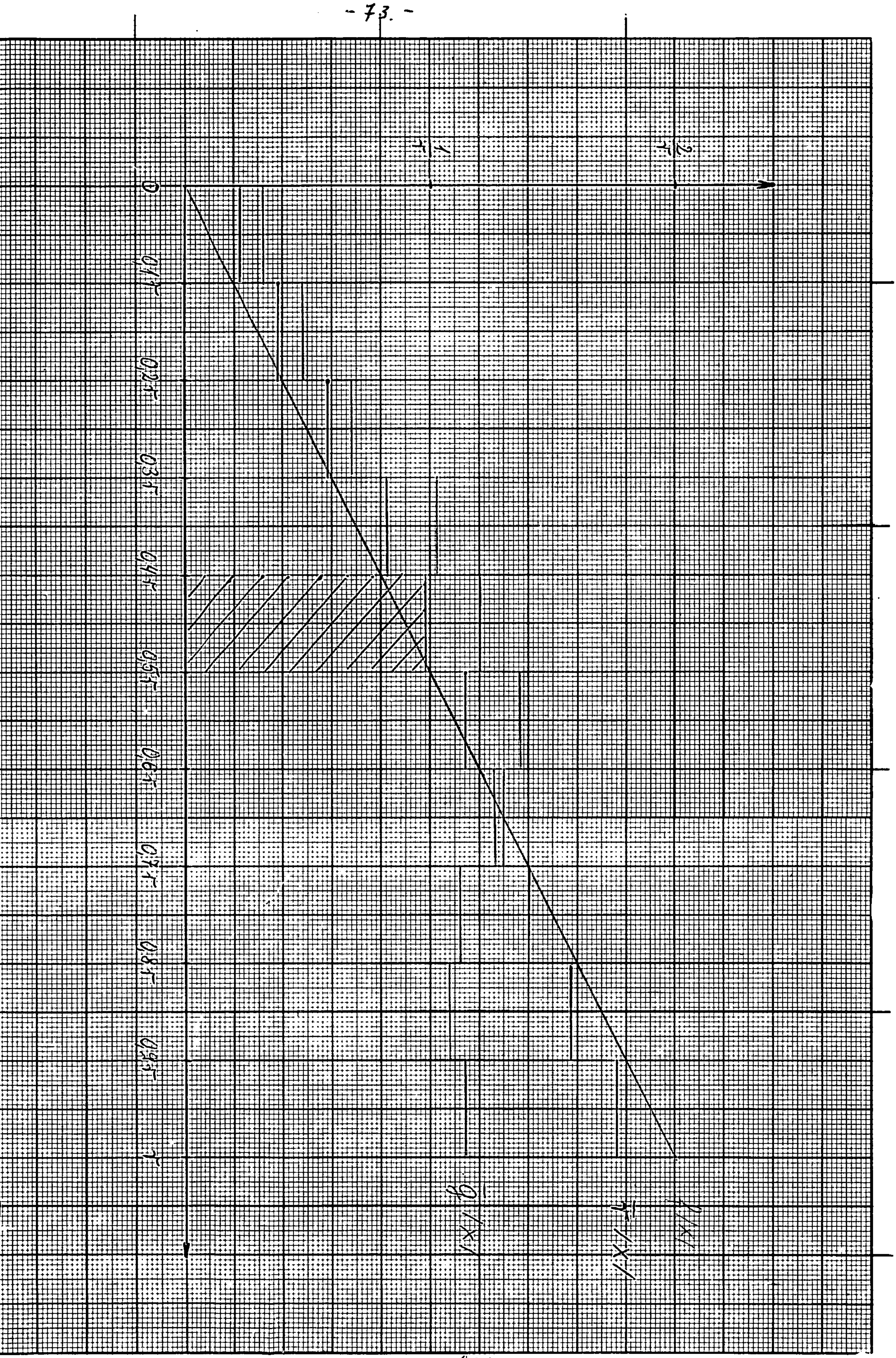
Ehhez az eredményhez egyszerü és meggyőződő matematikai uton az $F(x) = \frac{x^2 \pi}{r^2 \pi}$ eloszlásfüggvényből kiindulva is eljuthattunk volna.

A g(x) közelitését hasonló módon, de most már nem logikai feltevések, hanem kisérleti felmérések alapján végezzük. Ezt valóban végre is hajtottuk, a kisérleti eredményekből nyert empirikus szubjektiv sürüségfüggvény grafikonját a [16.] dolgozatunkból átvéve az 5. grafikonon ábrázoltuk kék szinnel.

Hasonló módon mérhetnénk fel r (x) függvényt közelitő empirikus szubjektiv sürüségfüggvényt is.



4. grafikon



5. grafikon

/A változás természetesen most is valamilyen I információ hatására ment végbe./ Az $\bar{r}(x)$ értékei fiktiv adatok, ezt a függvényt fekete vonallal ábrázoltuk.

A felvett információ értékét a három empirikus sürüségfüggvény alapján számoljuk. Az

$$\int_{0,4\,r}^{0,5\,r} \bar{r}(x)\,dx \quad \text{helyére például a zölddel}$$

besatirozott téglalap területét $/0,1\ r\ .\ 0,96\ .\ \frac{1}{r} = 0,096/$ irjuk.

$$\phi_\xi(I) \approx 0,01\ .\ \log \frac{0,022}{0,032} + 0,03 \log \frac{0,038}{0,048} +$$

$$+ \quad 0,05\ .\ \log \frac{0,054}{0,068} + 0,07 \log \frac{0,082}{0,14} +$$

$$+ \quad 0,09\ .\ \log \frac{0,096}{0,12} + 0,11 \log \frac{0,116}{0,136} +$$

$$+ \quad 0,13\ .\ \log \frac{0,135}{0,128} + 0,15 \log \frac{0,144}{0,112} +$$

$$+ \quad 0,17\ .\ \log \frac{0,156}{0,108} + 0,19 \log \frac{0,172}{0,116},$$

a számolásokat elvégezve

$$\phi_\xi(I) \approx 0,137.$$

E számolást követve nem szabad arra a következtetésre jutni, hogy nincs értelme a folytonos és diszkrét eset megkülönböztetésének.

Most valóban ugyanolyan módon történt a számolás, mint ahogyan diszkrét esetben tettük, látni kell azonban azt a lényeges különbséget, hogy az ott kapott információ értékek pontos eredmények voltak, a mostani eredményünk viszont csak közelitő érték.

Nem jelentéktelen a $\varnothing_{\xi}(I)$ meghatározására felirt zárt integrálalak sem, ugyanis ennek segitségével az előbbi példa számadataiból kiindulva sokkal pontosabban is meghatározhatnánk $\varnothing_{\xi}(I)$-t. Ezekből az adatokból ugyanis valamilyen approximációs módszerrel a $g(x)$ -et és az $r(x)$ -et igen jól közelitő folytonos függvény irható fel, amiből a fenti integrál alkalmazásával pontosabb eredményhez jutunk. Ilyen számolásra nem mutatunk be példát, de megjegyezzük, hogy a Lagrange-féle interpolációs polinom /ld. [12.] 368. old./ igen könnyen alkalmazható. Ennek előnye, hogy az integrálás sem ütközik komoly problémába, ha az $f(x)$ függvény sem tul nehezen kezelhető.

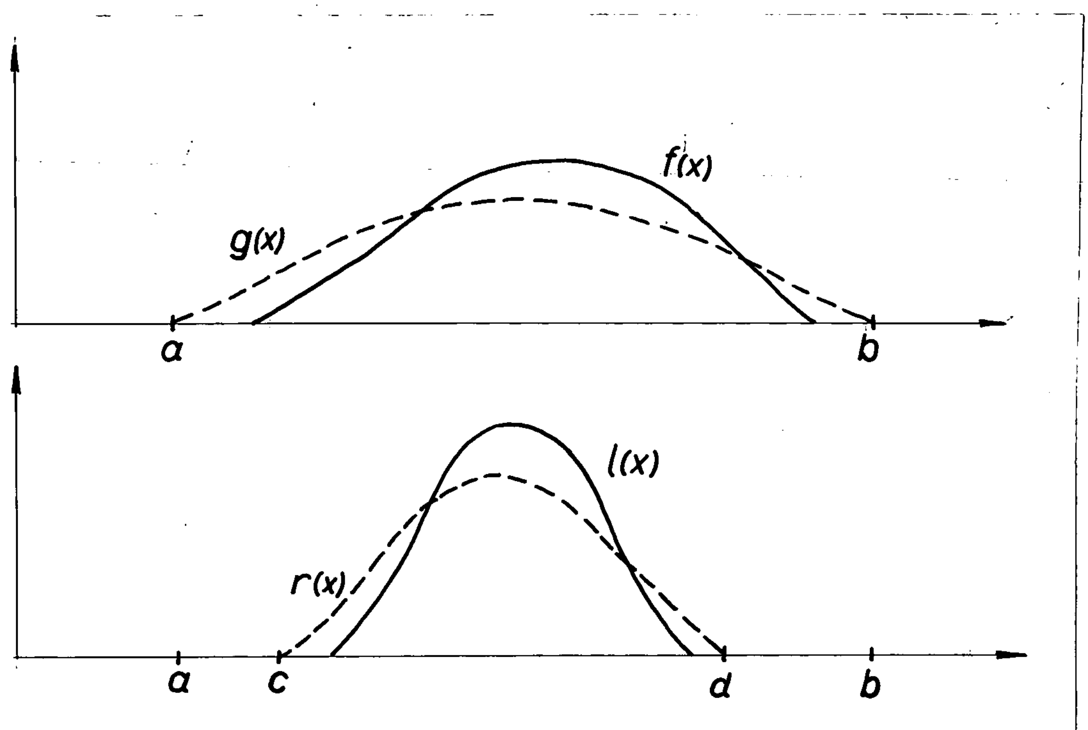
$\varnothing_{\xi}(I)$ értékének integrállal való kiszámitása akkor nélkülözhetetlen, ha az egyes sürüségfüggvények valamilyen zárt alakban felirható függvények, melyek jellegzetes eloszlástipushoz tartoznak. E kérdéskör ugyancsak a matematikai statisztikához tartozik.

Eddig a $P \times Q \times \mathcal{A} \xrightarrow{I} P \times Q' \times \mathcal{A}$ formulával leirható speciális esetet tárgyaltuk.

Térjünk át a legáltalánosabb, a $P \times Q \times \mathcal{A} \xrightarrow{I} P' \times Q' \times \mathcal{A}'$ formulával leirható esetre. Azt fogjuk megvizsgálni, hogy hogyan lehet mérni a szubjektum által felvett információt, ha a matematikai eloszlás is megváltozik, pontosabban az információ felvétele után a vizsgálati személynek, de főként a kisérletvezetőnek uj matematikai valószinüségi mezővel kell számolnia.

Jelöljük az eredeti matematikai sürüségfüggvényt $f(x)$-el, a szubjektivet $g(x)$ -el, az információ felvétele után a matematikai sürüségfüggvényt $h(x)$-el, a szubjektivet pedig $r(x)$ -el.

E függvényeket a 6. ábra szemlélteti.



6. ábra

Vegyünk fel olyan intervallumot, mely mind a négy függvény értelmezési tartományát tartalmazza. Az ábra szerint most az $[a;\ b]$ kielégiti e feltételt.

Osszuk fel az $[a;\ b]$ intervallumot n egyenlő részre, az osztáspontok: $a = x_0;\ x_1;\ \ldots\ ;\ x_n = b$.

A felvett információ értékét most is e két szubjektiv entrópia különbségeként számitjuk:

$$\Phi_\xi(I) \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)\, d\, x \;\log \frac{1}{\int_{x_i}^{x_{i+1}} g(x)\, d\, x} =$$

$$- \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} h(x) \;\log \frac{1}{\int_{x_i}^{x_{i+1}} r(x)\, d\, x}$$

Belátható, hogy $n \to \infty$ esetén a jobboldal konvergens, ennek határértéke adja $\Phi_\xi(I)$ pontos értékét, mégpedig

$$\Phi_\xi(I) = \int_a^b \left[f(x) \log \frac{1}{g(x)} - h(x) \cdot \log \frac{1}{r(x)} \right] d\, x \,.$$

Bizonyitás vázlata: Az előbbi összeg

$$\sum_{i=0}^{n-1} f(z_i)\left(x_{i+1} - x_i\right) \log \frac{1}{g(z_i')\,\left(x_{i+1} - x_i\right)} -$$

$$- \sum_{i=0}^{n-1} h(s_i)\left(x_{i+1} - x_i\right) \log \frac{1}{r(s_i')\left(x_{i+1} - x_i\right)}$$

alakban irható, ahol z_i ; z_i' ; s_i ; s_i' alkalmasan választott x_i és x_{i+1} közé eső pontok.

Vonjuk össze a két összeget és $(x_{i+1} - x_i)$ -t emeljük ki, majd hajtsunk végre egy olyan átcsoportositást, hogy a $\log \frac{1}{x_{i+1} - x_i}$ -t is kiemelhessük két tagból. Végül a következő összeghez jutunk:

$$\sum_{i=o}^{n-1} \left[f(z_i) - h(s_i)\right] (x_{i+1} - x_i) \log \frac{1}{x_{i+1} - x_i} +$$

$$+ \sum_{i=o}^{n-1} \left[f(z_i) \log \frac{1}{g(z_i')} - h(s_i) \log \frac{1}{r(s_i')}\right](x_{i+1} - x_i)$$

Mivel $x_{i+1} - x_i = \frac{b-a}{n}$, ezért az első összeg

$$\sum_{i=o}^{n-1} \left[\int_{x_i}^{x_{i+1}} f(x)\, d\, x - \int_{x_i}^{x_{i+1}} h(x) d\, x\right] . \log \frac{n}{b-a}$$

alakban irható. Ebből $\log \frac{n}{b-a}$ -t kiemelve /ez nem függ i-től/

$$\log \frac{n}{b-a} \left[\int_a^b f(x)\, d\, x - \int_a^b h(x)\, d\, x\right]$$ kifejezéshez jutunk,

ami azonosan 0, mert

$$\int_a^b f(x)\, d\, x = \int_a^b h(x)\, d\, x = 1.$$

Az átalakitott összeg másik tagja pedig éppen a $\Phi_\xi(I)$ -re felirt integrál közelitő összege.

A

$$\varnothing_{\xi}(I) = \int_a^b \left[f(x) \log \frac{1}{g(x)} - h(x) \log \frac{1}{r(x)} \right] dx$$

eredményhez most is eljuthattunk volna ugy is, ha az

$$\int_a^b f(x) \log \frac{1}{g(x)} dx \quad \text{és} \quad \int_a^b h(x) \log \frac{1}{r(x)} dx$$

integrálok, mint szubjektiv entrópiák különbségét vesszük.

Ha $f(x) = h(x)$, akkor a már ismert

$\int_a^b f(x) \log \frac{r(x)}{g(x)} dx$ adódik speciális esetként, tehát

valóban általánosabb definicióhoz jutottunk.

Most is igaz az, hogy ha az integráljel alatti függvényeknek csak az empirikus alakját ismerjük, akkor a legelőször felirt összeg alapján kell számolnunk a felvett információ értékét.

Mivel még a zárt integrálalak használatára nem mutattunk példát, ezért most egy fiktiv kisérlet elemzése utján ezt tesszük, annak ellenére, hogy a számolás igy kissé mélyebb matematikai módszereket igényel. Viszont meglátszik majd e módszerben rejlő előny, hogy az eddigieknél sokkal kevesebb mérés alapján is lehet következtetni /persze bizonyos feltételek elfogadásával/ a felvett információ nagyságára.

A vizsgált személytől azt kérdezzük, hogy milyen hosszu egy asztal. E kérdésre nem adhat határozott választ, hanem /természetesen a szó szoros értelmében nem tudatosan/ egy szubjektiv valószinüségi mező alakul ki benne, amit akkor tudnánk jól jellemezni, ha a szubjektiv sürüségfüggvényt ismernénk.

Ennek részletes mérések alapján való közelitésére, az empirikus szubjektiv sürüségfüggvény felrajzolására már mutattunk példát. Most feltesszük, hogy ilyen részletes mérésre nincs lehetőségünk. Mindössze a következő három kérdést tesszük fel a v.sz.-nek: Szerinte milyen hosszu lehet a legkisebb asztal? /A játékasztalokat leszámitva./ Milyen hosszu egy "átlagos méretü" asztal? Milyen hosszu az a legnagyobb asztal, amit el tud képzelni, hogy gyártanak? /Itt valódi használhatóságra gondoljon./

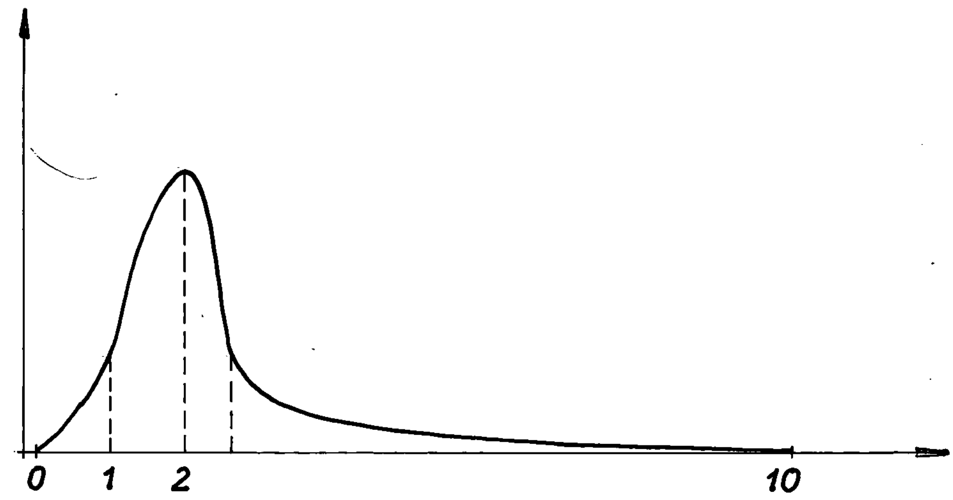
Tegyük fel, hogy rendre a következő válaszokat kapjuk: 0,2 m; 2 m; 10 m.

E három adat nem határozza meg egyértelmüen a szubjektiv sürüségfüggvényt, ezek alapján még akárhányat fel lehetne irni, tehát további feltevésekkel kell élnünk.

Ami a felmérésből következik az, hogy a keresett $g(x)$ értelmezési tartománya /pontosabban, ahol $g(x) \neq 0$ / a $[0,2\ ;\ 10]$ intervallum és a 2-nél veszi fel a maximumát, azaz egy ferde eloszlásról van szó.

Ezután már a felmérőn mulik, hogy milyen feltételeket szab még ki a szubjektiv sürüségfüggvényre.

Mi a következő megkötéseket tesszük: $g(x)$ 0,2-től 1-ig először lassan, majd gyorsabban nő, 1-től 2-ig ugyancsak nő, ott felveszi a maximumát, onnan először gyorsan, majd egyre lassabban csökken 10-ig. /ld. 7. ábra/.



7.ábra

Elképzelhető, hogy egy közismert eloszlással irjuk le a szubjektiv eloszlást, mi azonban most arra is törekszünk, hogy a sürüségfüggvényeket polinom alakban irjuk fel.

A $g(x)$ -et /majd a többi sürüségfüggvényt is/ a következő alakban fogjuk keresni:

$$g(x)=\begin{cases} 0 & \text{, ha } x \leq 0{,}2 \\ \dfrac{(x-0{,}2)^2}{a} & \text{, ha } 0{,}2 \leq x \leq 1 \\ \dfrac{(x-0{,}2)^2 - b(x-1)^2}{a} & \text{, ha } 1 \leq x \leq 2{,}5 \\ \dfrac{(10-x)^2}{c} & \text{, ha } 2{,}5 \leq x \leq 10 \\ 0 & \text{, ha } x \geq 10 \end{cases}$$

A képletekben szereplő a; b; c betük paraméterek, melyek a folytonos sürüségfüggvényre jellemző és a kisérletvezető által meghatározott tulajdonságokból határozhatók meg. A második képletben szereplő x-1 kifejezés biztositja, hogy az első és a második parabolának az 1-es pontban közös pontjuk legyen.

A második és a harmadik parabola találkozási pontját /az értelmezési tartományok megadásakor/ a 2,5-nél jelöltük ki. /Ez a kijelölés bizonyos mértékig önkényes, a kisérletvezető határozza meg, mi a már leirt feltételeinket tartjuk szem előtt./

Mivel a v.sz. a 2 m-es hosszuságot tartotta "legvalószinübbnek", ezért a második parabolát célszerü ugy alakitani, hogy maximuma a 2-nél legyen. Ez akkor következik be, ha az 1-nél felvett értéke megegyezik a 3-nál felvett értékével /igy az x=2 egyenesre szimmetrikus lesz/.

Ez természetesen nem azt jelenti, hogy a szóbanlevő valószinüségi változó várható értéke 2 lesz. Ezt a feltételt is lehetne teljesiteni az általunk bemutatott módszerhez hasonló okoskodással.

Az utóbbi feltételből b határozható meg:

$$\frac{(1-0{,}2)^2}{a} = \frac{(3-0{,}2)^2 - b(3-1)^2}{a},$$

$$0{,}64 = 7{,}84 - 4\,b\,, \quad \text{amiből} \quad b = 1{,}8$$

Második feltételünk, hogy a második és a harmadik parabola a 2,5 abszcisszáju pontban metsze egymást. Ez a következő egyenletet adja:

$$\frac{(2{,}5-0{,}2)^2 - 1{,}8\cdot(2{,}5-1)^2}{a} = \frac{(10-2{,}5)^2}{c},$$

$$\frac{1{,}24}{a} = \frac{56{,}25}{c}\,, \quad \text{ebből}$$

$$c = 45{,}36\ a.$$

Már csak egy független paraméterünk marad, amit ugy kell megválasztanunk, hogy az

$$\int_{0,2}^{10} g(x)\,d\,x = 1 \qquad \text{feltétel teljesüljön,}$$

ami a következő egyenlethez vezet:

$$\frac{1}{a}\int_{0,2}^{1}(x-0,2)^2\,d\,x + \frac{1}{a}\int_{1}^{2,5}\left[(x-0,2)^2 - 1,8(x-1)^2\right]d\,x +$$

$$+ \frac{1}{45,36\,a}\int_{2,5}^{10}(10-x)^2\,d\,x = 1$$

Számitsuk ki az egyes integrálokat!

$$\int_{0,2}^{1}(x-0,2)^2\,d\,x = \left[\frac{x^3}{3} - 0,2\,x^2 + 0,04\,x\right]_{0,2}^{1} = 0,171\ ,$$

$$\int_{1}^{2,5}\left(-0,8\,x^2 + 3,2\,x - 1,76\right)d\,x = \left[-\frac{0,8\,x^3}{3} + 1,6\,x^2 - 1,76\,x\right]_{1}^{2,5} =$$

$$= 1,86\ ,$$

$$\int_{2,5}^{10}\left(x^2 - 20\,x + 100\right)d\,x = \left[\frac{x^3}{3} - 10\,x^2 + 100\,x\right]_{2,5}^{10} = 140,6\ .$$

Az előbbi egyenletünk tehát a

$$\frac{0,171}{a} + \frac{1,86}{a} + \frac{140,6}{45,36a} = 1 \qquad \text{alakot ölti, ahonnan } a=5,13\ .$$

A keresett függvény tehát:

$$g(x)=\begin{cases}\dfrac{(x-0,2)^2}{5,13} & \text{, ha } 0,2\le x\le 1\\[2ex] \dfrac{(x-0,2)^2-1,8(x-1)^2}{5,13} & \text{, ha } 1\le x\le 2,5\\[2ex] \dfrac{(10-x)^2}{232,7} & \text{, ha } 2,5\le x\le 10\end{cases}$$

Polinom alakban irva:

$$g(x)=\begin{cases}0,195\ x^2-0,078\ x+0,0078 & \text{, ha } 0,2\le x\le 1\\ -0,156\ x^2+0,624\ x-0,343 & \text{, ha } 1\le x\le 2,5\\ 0,0043\ x^2-0,086\ x+0,43 & \text{, ha } 2,5\le x\le 10\end{cases}$$

Az $f(x)$ matematikai sürüségfüggvény meghatározása szóról-szóra megegyezik a fentiekkel, az adatok meghatározása most nem kérdések, hanem mérések alapján történik. A kisérletvezető megmér néhány jellegzetes asztalt és tegyük fel, hogy azt tapasztalja, hogy 0,5 m-nél rövidebb asztal nincs /vagy az ennél rövidebbeket egyszerüen nem fogadja el asztalnak/, a mindennapi életünkben előforduló asztalok hosszának átlaga 1,8 m, 5 m-nél hosszabb asztal szintén nincs /a speciális célokra készült asztalokat nem veszi figyelembe/.

E tapasztalatok alapján $f(x)$ értelmezési tartománya $[0,5\ ;\ 5]$, a 2 m szerepét most az 1,8 m veszi át. A csatlakozási pontoknak pedig az 1-et és a 2-t választjuk, igy az $f(x)$ függvényt a következő alakban keressük:

$$f(x)=\begin{cases}\dfrac{(x-0{,}5)^2}{a} & \text{, ha } 0{,}5\le x\le 1\\[2mm] \dfrac{(x-0{,}5)^2-b(x-1)^2}{a} & \text{, ha } 1\le x\le 2\\[2mm] \dfrac{(5-x)^2}{c} & \text{, ha } 2\le x\le 5\end{cases}$$

A szükséges számitásokat elvégezve az

$$f(x)=\begin{cases}0{,}826\,x^2-0{,}826\,x+0{,}207 & \text{, ha } 0{,}5\le x\le 1\\ -0{,}517\,x^2+1{,}86\,x-1{,}136 & \text{, ha } 1\le x\le 2\\ 0{,}057\,x^2-0{,}57\,x+1{,}435 & \text{, ha } 2\le x\le 5\end{cases}$$

eredményhez jutunk.

Ha közöljük a v.sz.-lyel, hogy konyhai asztalról van szó, akkor e közlés mint információ szerepel számára, hisz lényegesen megváltozik az asztal hosszáról alkotott véleménye, a mi elképzelésünk szerint egy uj szubjektiv sürüségfüggvény lép fel.

Tegyük fel, hogy a szubjektiv sürüségfüggvényre a kisérletvezető /kérdései alapján/ a következő jellegzetes pontokat határozza meg:

1 m a legrövidebb, 1,8 m az átlag, 3 m a leghosszabb asztal. A csatlakozási pontokat 1,4-nél és 2-nél jelöli ki. $r(x)$ -et a következő alakban állitjuk elő:

$$r(x)=\begin{cases}\dfrac{(x-1)^2}{a} & \text{, ha } 1\le x\le 1{,}4\\[2mm] \dfrac{(x-1)^2-b(x-1{,}4)^2}{a} & \text{, ha } 1{,}4\le x\le 2\\[2mm] \dfrac{(3-x)^2}{c} & \text{, ha } 2\le x\le 3\end{cases}$$

A paraméterek meghatározása után:

$$r(x)=\begin{cases} 3,53\ x^2 - 7,07\ x + 3,53 & ,\ \text{ha}\quad 1 \le x \le 1,4 \\ -\ 3,53\ x^2 + 12,72\ x - 10,32 & ,\ \text{ha}\quad 1,4 \le x \le 2 \\ 0,99\ x^2 - 5,94\ x + 8,91 & ,\ \text{ha}\quad 2 \le x \le 3 \end{cases}$$

Végül tegyük fel, hogy a kisérletvezető e valószinüségi változóra /mérések alapján/ a következő adatokat tartja reálisnak: legrövidebb 1,2 m, átlag 1,6 m, leghosszabb 2,5 m. A csatlakozási pontokat pedig 1,4-ben és 1,8-ban jelöli meg. A felsorolt adatok alapján az információfelvétel utáni matematikai sürüségfüggvényt a következő képletek irják le:

$$l(x)=\begin{cases} 25,64\ x^2 - 61,54\ x + 36,92 & ,\ \text{ha}\quad 1,2 \le x \le 1,4 \\ -\ 25,64\ x^2 + 82,05\ x - 63,59 & ,\ \text{ha}\quad 1,4 \le x \le 1,8 \\ 2,08\ x^2 - 10,39\ x + 13 & ,\ \text{ha}\quad 1,8 \le x \le 2,5 \end{cases}$$

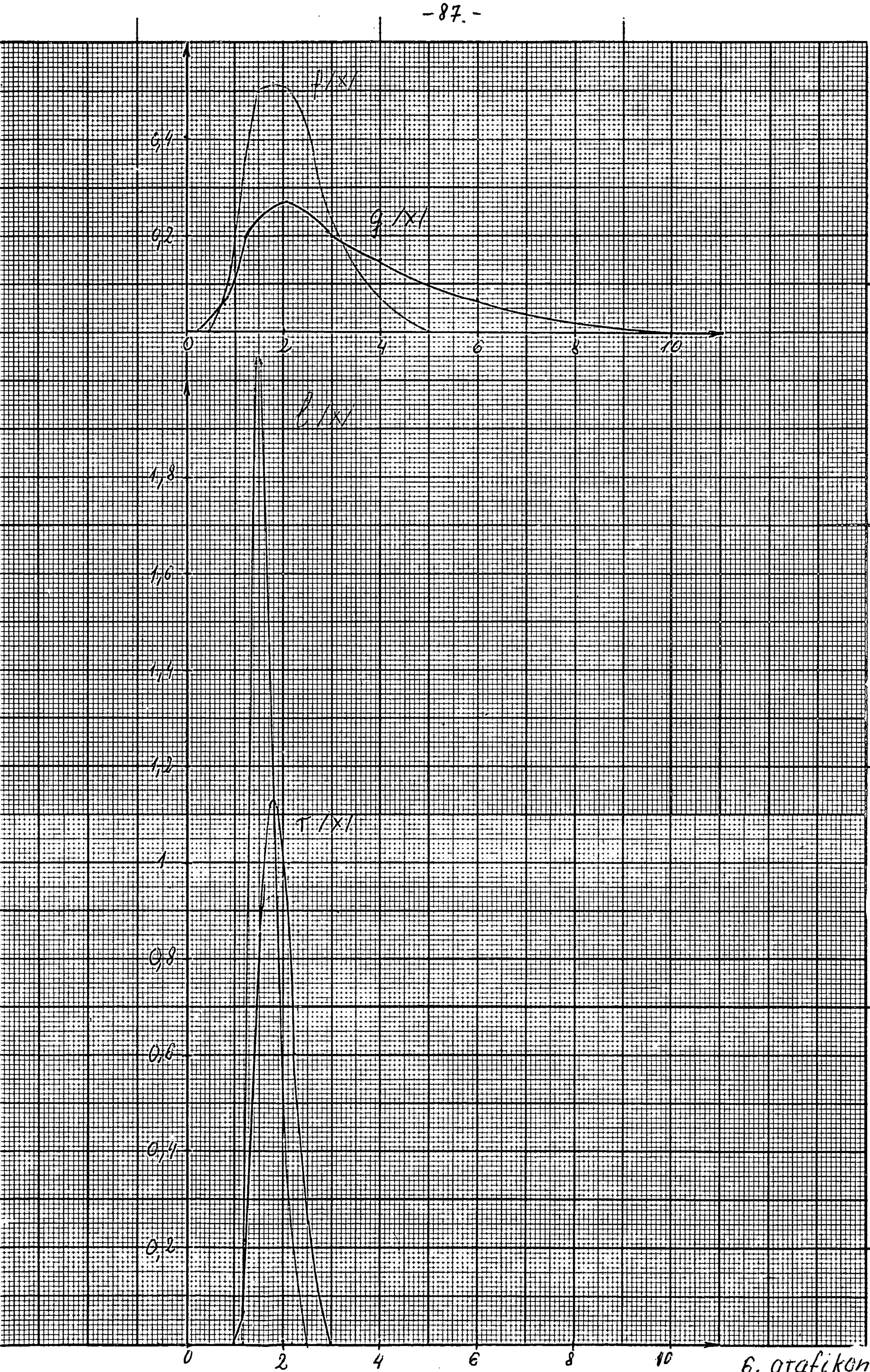
A felirt négy sürüségfüggvényt a 6. grafikonon ábrázoltuk. Piros vonallal a matematikai, kékkel pedig a szubjektiv sürüségfüggvények grafikonját huztuk meg.

Most már használhatjuk az információ mérésére levezetett

$$\Phi_{\xi}(I) = \int_a^b \left[f(x) \log \frac{1}{g(x)} - h(x) \log \frac{1}{r(x)} \right] d\,x$$

integrált.

Bár a legbővebb intervallum a $[0,2;\ 10]$, az integrálást mégis elegendő a $[0,5\ ;\ 5]$ intervallumon végezni, mert ezen kivül csak a $g(x)$ nem nulla. Az integrál kiszámitásánál egy kis nehézséget okoz, hogy az egyes függvények három-három képlettel vannak megadva, melyek ráadásul különböző intervallumokon értelmezik az egyes függvényeket.



6. grafikon

Fel kell tehát bontanunk a $[0,5 ; 5]$-ot olyan részintervallumokra, melyeken belül minden sürüségfüggvényt csak egy polinom ir le. A keresett integrált igy több integrál összegeként kapjuk:

$$\int_{0,5}^{5} = \int_{0,5}^{1} + \int_{1}^{1,4} + \int_{1,4}^{1,8} + \int_{1,8}^{2} + \int_{2}^{2,5} + \int_{2,5}^{5}$$

A számolás nem nehéz /bár a középiskolainál kissé magasabb szintü jártasságot igényel/, de hosszadalmas, ezért nem irjuk le.

Megjegyezzük, hogy az integrálás müvelete kikerülhető, ha milliméterpapiron ábrázoljuk az

$$x \longrightarrow f(x) \cdot \log \frac{1}{g(x)} - l(x) \cdot \log \frac{1}{r(x)}$$

függvényt, és az alatta levő területet megfelelő pontossággal becsüljük.

Az információ értékének meghatározása talán tul hosszadalmasnak és nehézkesnek tünhet. Észre kell azonban venni, hogy az egyes eljárások ismétlődtek, tehát programozhatók, a számitógép használata tehát több kisérlet elemzésekor kifizetődő.

A folytonos valószinüségi változóval kapcsolatban még több problémát lehetne felvetni. Ugy gondoljuk azonban, hogy elgondolásunk lényegét sikerült bemutatni, az ezen tul felmerülő problémák most már inkább csak matematikai jellegüek.

Nem térhetünk ki azonban az elől a pszichológiai szempontból igen fontos speciális eset vizsgálata elől, amikor a valószinüségi változó csak egy értéket vehet fel. Ez a helyzet áll elő, ha például egy konkrét rud hosszát becsültetjük a kisérleti személlyel.

A rud hossza határozott érték, a kisérleti személy viszont csak "valószinüsiteni" tudja a rud hosszát. A kisérleti személyben tehát kialakul egy szubjektiv sürüségfüggvény, ahol valószinüségi változó a rud hossza. Jelöljük e szubjektiv sürüségfüggvényt most is $g(x)$ -el.

Ha a rud valódi hossza h, akkor annak matematikai valószinüsége, hogy a rud hossza h-nál kisebb 0, annak matematikai valószinüsége, hogy a rud hossza kisebb egy h-nál nagyobb x számnál 1. Ezt az eloszlást a következő eloszlásfüggvény irja le.

$$F(x) = \begin{cases} 1 & , \text{ ha } \quad x \geq h \\ 0 & , \text{ ha } \quad x < h \end{cases}$$

Vegyünk fel a h körül egy $[h - \Delta x;\ h + \Delta x]$ intervallumot. Ha a valódi hosszuságot közöljük a kisérleti személlyel, akkor az adott intervallumra vonatkozóan

$$\log \frac{1}{\int_{h-\Delta x}^{h+\Delta x} g(x)\, dx}$$ meglepetést okozunk.

Ez egyuttal a szubjektiv entrópia is, hisz annak matematikai valószinüsége, hogy a valószinüségi változó ebbe az intervallumba esik 1.

Ha a szubjektiv eloszlás valamilyen I információ hatására /például e rud mellé teszünk a kisérleti személy által ismert hosszuságu rudat/ $g(x)$ -ről $r(x)$ -re változik, akkor azt mondhatjuk, hogy erre az intervallumra vonatkozóan

$$\log \frac{1}{\int_{h-\Delta x}^{h+\Delta x} g(x)\, dx} - \log \frac{1}{\int_{h-\Delta x}^{h+\Delta x} r(x)\, dx}$$

nagyságu információt közöltünk.

Ha az intervallum hosszát egyre kisebbre változtatjuk, vagyis ha $\Delta x \to 0$, akkor e különbség

$$\log \frac{r(h)}{g(h)}$$ értékhez tart.

Tehát:

$$\phi_h(I) = \log \frac{r(h)}{g(h)} .$$

Figyelemre méltó, hogy az információ értéke most már nem függ semmilyen intervallum hosszától, csupán a szubjektiv sürüségfüggvények h-beli értékétől.

Ez megfelel annak a diszkrét esetnél tárgyalt információfelvétel mérésének, amikor a válasz csak egy meghatározott A esemény bekövetkeztének közlésekor lehet helyes.

Ezzel az első és egyben alapproblémánkat megoldottuk. Egy $\mathcal{A}$ eseményrendszerre /ami lehet folytonos valószinüségi változóval leirható/, pontosabban egy $P \times \mathcal{A}$ valószinüségi mezőre nézve definiáltuk az információ és annak értékének fogalmát. Az információ értékének meghatározására több példát mutattunk. Mielőtt azonban tovább általánositanánk, egy irodalomból vett példán szeretnénk még nyomatékot adni módszerünk elméleti jelentőségének és gyakorlati pontosságának.

IV. Egy tárgy hosszának észlelésében tartalmazott információ értéke.

Az irodalomban /ld. [10.] / kedvelt alapkérdésként szerepel annak mérése, hogy egy tárgy /pl. rud/ hosszának megfigyelése mekkora információt hordoz a k.sz. számára. Az előző fejezet példája is hasonló kisérletet elemez, itt azonban az a célunk, hogy e kérdést részletesebben megvizsgáljuk, bemutatva a felületesebb és a mélyebb meggondolások közti különbséget, egyben abban is reménykedünk, hogy sikerül módszerünk előnyét is szemléltetni.

Ha a k.sz.-lyel egy rud hosszát becsültetjük meg, a mi értelmezésünk szerint nincs értelme információfelvételről beszélni, ha nem volt kezdeti bizonytalanság. Más a helyzet, ha például a bemutatás előtt azt mondjuk a kisérleti személynek, hogy "egy horgászbotot fog látni, becsülje meg ennek hosszát". Ebben az esetben már a bemutatás előtt fellép a kezdeti bizonytalanság. A bot bemutatása ezt a bizonytalanságot csökkenti, igy információfelvétel történik, aminek értékét mérni tudjuk előbbi módszerünkkel. Ha az előbbi megjegyzést nem tesszük, akkor is kap információt a kisérleti személy, de ennek mérése nem lesz addig egyértelmü, amig viszonyitási alapot nem adunk, ami a mi megfogalmazásunk szerint az $\mathcal{A}$ eseményrendszer kijelölését jelenti.

Egy hosszuság észlelésekor felvett információ mérésének ismert speciális esete, amikor egy adott intervallumon megjelenő pont helyzetét kell becsülnie a kisérleti személynek. Ezzel egyenértékü, csak más megfogalmazásu az a kisérlet, amikor egy müszer mutatójának helyzetét kell megbecsülni olyan esetben, amikor a müszer lapjának nincs beosztása.

A kisérlet konkrét végrehajtása például a következő: A kisérleti személytől bizonyos távolságra levő, adott s hosszuságu szakaszon véletlenszerüen kijelölnek egy pontot, majd a kisérleti személy kezében levő, ugyanolyan hosszuságu, valahány egyenlő részre osztott szakaszon a ksz.-nek be kell jelölnie, hogy szerinte mely részintervallumba esik a pont, tehát az adott skála alapján a pont helyzetét kell megállapitania.

Kisérletileg kimutatták /G. A. Miller "Seven plus or minus twO" törvénye/, hogy a kisérleti személy akkor tud biztos választ adni, ha maximálisan 9 /pontosabban 7 ± 2/ egyenlő részre osztjuk a kisérleti személy kezében levő szakaszt. Ez alapján következtetnek arra, hogy ilyen uton a szubjektummal log 9 bit információ közölhető. Rámutatunk e következtetés labilisságára.

E megállapitást a mi módszerünk szerint ugy magyarázhatjuk, hogy ha az egész intervallumot 9 egyenlő részre osztjuk, akkor a pont egyenlő valószinüséggel kerülhet bármely részintervallumba, tehát $\frac{1}{9}$ annak a valószinüsége, hogy egy adott intervallumba esik.

A matematikában használatos képlet alapján

$$\sum_{i=1}^{9} \frac{1}{9} \log \frac{1}{\frac{1}{9}} = \log 9$$

az eredeti entrópia és /mivel a bemutatás után már tudja, hogy a kilenc közül melyikbe esett/ log 1 = 0 lesz az uj entrópia, a kettő különbsége pedig a felvett információ értékét, log 9-et adja.

Az első, amit észrevehetünk, hogy itt feltételezik a szubjektiv és matematikai valószinüség megegyezését, pontosabban nincs szubjektiv valószinüség.

Bizonyos feltételek mellett ugyanerre az eredményre jutunk, ha a kisérleti személynél levő szakaszt nem 9, hanem $n = 9 \cdot m$ /az egyszerüség kedvéért m természetesszám/ egyenlő részre osztjuk. Ahhoz, hogy a kisérleti személy tudomására hozzuk, hogy a pont melyik intervallumba került

$$\sum_{i=1}^{9m} \frac{1}{9m} \log \frac{1}{\frac{1}{9m}} = \log 9\,m$$ információra van szükség.

/Természetesen itt még mindig az elterjedt számolási módot alkalmazzuk./ A pont puszta megfigyelésével még nem kapja meg ezt az információértéket, ha $m > 1$, mert az előbbi törvény szerint csak azt tudja eldönteni a kisérleti személy, hogy mely m darab részintervallumot tartalmazó intervallumban helyezkedik el a pont. Ha ezen belül egyenletes eloszlást tételezünk fel, akkor log m információra van még szükség ahhoz, hogy a pont helyét az adott beosztásra vonatkozóan pontosan megtudja.

A felvett információ tehát most is

$$\log 9\,m - \log m = \log 9 \quad /bit/.$$

Ezt a felbontást azért végeztük el, hogy világosabb legyen milyen feltételekkel kell élnünk, ha e kisérletből azt szürjük le, hogy ez az észlelés log 9 bit információt hordoz.

Mindenekelőtt fel kellett tételeznünk, hogy a kezdeti szubjektiv eloszlás egyenletes. Már ebben sem lehetünk biztosak, meglehet, hogy a szubjektum előtt nincs olyan esélye a szakasz két szélének, mint a közepének. Itt elég a szubjektiv eloszlás felmérésénél tapasztalt /ld. [16] / "széli effektusra" gondolni. A matematikai eloszlás egyenletességét nem vitatjuk, ezt a kisérletvezető, illetve a feltételek határozzák meg. Hallgatólagos feltétel volt az is, hogy a pont helyzetétől függetlenül $\frac{1}{9}$ "pontosságig" képes egyértelmüen eldönteni a pont helyét a kisérleti személy. Valószinünek tartjuk, hogy a teljes szakasz végpontjainak közelébe eső pont helyzetét ennél jóval pontosabban tudja megállapitani.

Legnagyobb problémát azonban abban a /szintén hallgatólagos/ feltételben látjuk, hogy az adott $\frac{1}{9}$ részen belül már egyenletesnek tételezi fel ez a felfogás a szubjektiv valószinüségeloszlást. /Pontosabban, mivel ott szubjektiv eloszlásról szó sincs, akkor jutunk ehhez az eredményhez, ha egyenletes szubjektiv eloszlással számolunk./

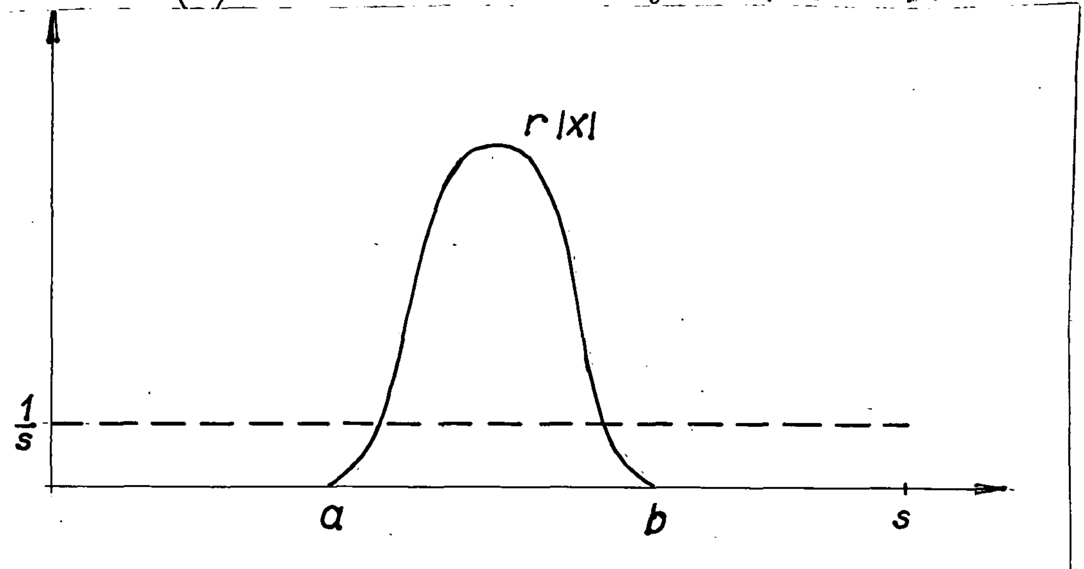
A mi felfogásunk szerint ennél a kisérletnél is fel kell mérni a szubjektiv sürüségfüggvényeket.

Legyen $f(x) = \frac{1}{s}$ a kezdeti matematikai sürüségfüggvény /s a teljes intervallum hossza, tehát egyenletes eloszlást tételeztünk fel/.

Vegyük $g(x) = \frac{1}{s}$-nek a kezdeti szubjektiv sürüségfüggvényt is. Ez utóbbit valószinüleg cáfolná egy konkrét felmérés, de kisérleti adatok hiányában fogadjuk el ezt a feltételt, ez nem zavarja mondanivalónk lényegét, hisz a nagyobb problémát nem ebben a feltevésben látjuk.

A kisérlet végrehajtása után matematikai sürüségfüggvény nem lesz, mert a pont helyzete már rögzitetté válik. Az uj szubjektiv sürüségfüggvényt jelöljük $r(x)$-el. Az előbb emlitett törvény szerint ennek értelmezési tartománya nem lehet nagyobb $\frac{s}{9}$-nél, mert ekkor-a a maximális intervallum, melybe biztosan bejelöli a kisérleti személy a pont helyét.

Tegyük fel, hogy a pont a $[0; s]$ intervallum 0 pontjától h távolságra esett. Megint fenti törvényből következik, hogy h eleme $r(x)$ értelmezési tartományának /8.ábra/.



8. ábra

A felvett információ értékét a már leirt módon számolhatjuk ki:

$$\emptyset_h(I) = \log \frac{r/h/}{\frac{1}{s}} = \log s \cdot r/h/ \ .$$

Ha most az előbbiekkel való összehasonlitás céljából feltesszük, hogy $\frac{s}{9} = b - a$ /azaz r /x/ értelmezési tartománya éppen $\frac{s}{9}$ hosszu/, akkor

$$\phi_h(I) = \log 9 \,/\, b - a \,/\, r\,/h/ .$$

Ha r /x/ egyenletes eloszlás sürüségfüggvénye, akkor /b-a/r /h/ = 1, tehát a felvett információ értéke valóban log 9. Ha viszont r/x/ nem egyenletes eloszlást ir le, akkor /b-a/, r /h/ 1-nél nagyobb, de 1 -nél kisebb is lehet, ami maga után vonja, hogy a felvett információ log 9-nél nagyobb, vagy kisebb is lehet. Véleményünk szerint ebben az esetben a felvett információ értéke általában nagyobb log 9-nél, hisz valószinü, hogy r/x/ maximum helye h-hoz közel esik.

Természetesen az általunk használt felmérésnél is részintervallumokra kell osztani az s távolságot /sőt 9-nél több egyenlő részre/, de a lényeges eltérés ott mutatkozik, hogy mi nem azt kérdezzük a v.sz.-től, hogy szerinte melyik intervallumba esik a pont, hanem annak szubjektiv valószinüségét mérjük fel, hogy az egyes részintervallumokba beleesik a pont.

Bár itt ugy tünik, csak pszichológiai fejtegetésekről van szó, mégsem sok kellene hozzá, hogy pedagógiai témára térjünk át. Ezzel a példával megint alátámaszthatnánk azt a /már emlitett/ véleményünket, hogy a feleletválasztásos kérdéseket /teszteket/ ugy kellene összeállitani, hogy ott ne csak egy válasz megjelölésére legyen lehetőség, hanem a több megadott választ valószinüsiteni tudja a tanuló. Igy ott is reálisabb képet kapnánk.

Térjünk vissza a pszichológiai kisérlet birálatához.

Nem valószinü az sem, hogy a felvett információ értéke s-től független. Érdekes lenne például megvizsgálni, hogy hogyan változna $g/x/$, illetve a felvett információ értéke, ha s-t ugy növelnénk, hogy közben $\frac{h}{s}$ állandó maradna.

Félreértés ne essék, előbbi fejtegetéseinkkel nem az emlitett törvényt biráltuk /azt bizonyosan több soron ellenőrizték/, csupán azt, hogy abból nem következik, hogy egy ilyen észlelés log 9 bit információt jelent.

A birált felmérési mód egy sulyos ellentmondáshoz is vezethet.

Vegyünk megint egy s hosszuságu szakaszt, ezen fogjuk kijelölni a pontot. A kisérleti személy kezében két ugyanilyen szakasz van, az egyik 5, a másik 7 egyenlő részre van osztva. Az első esetben a szubjektiv entrópia log 5, a másodiknál pedig log 7. Előfordulhat, hogy a kijelölt pont valamely ötös beosztáshoz igen közel esik, igy a kisérleti személy nem tudja eldönteni az 5-ös beosztásu szakaszon, hogy melyiket jelölje meg, a 7-es beosztásun viszont határozott helyes választ tud adni. Ezek szerint az észlelés utján nem kapott log 5 információértéket a kisérleti személy, de log 7 értéket igen? Persze ez abszurdum, az általunk javasolt módszernél ez nem fordulhat elő. A mi felfogásunk szerint ennél a kisérletnél két eseményrendszer van, az elsőre vonatkozóan valóban log 5-nél kisebb információt kapott, a kisérleti személy, a másodikra vonatkozóan pedig log 7-t. Ez csak alátámasztja azt a véleményünket, hogy egy közlés információtartalmát nem lehet eseményrendszertől függetlenül vizsgálni.

Végül két példát mutatunk egy η valószinüségi változóra vonatkozó információ értékének közvetett meghatározására. E példák bemutatását az ilyen tipusu helyzetek gyakorisága indokolja.

Egy terület becslése kétféleképpen történhet. Az egyik mód, amikor az adott területet mint egészet figyeljük és számitás nélkül becsültetjük meg a terület mértékét. Ebben az esetben a területet tekinthetjük valószinüségi változónak, a megfigyeléssel szerzett információ értékét a fentiek alapján számithatjuk. Másik lehetőség, hogy az adott terület meghatározó adatait /bizonyos hosszuságokat/ becsüljük, és e becsült értékekből számitjuk a területet. Általában az utóbbit tesszük "nevezetes" sikidomok /háromszög, kör, téglalap, stb/ területének becslésénél. Ez az eljárás az ember számára általában természetesebb, mert nagyobb gyakorlata van a hosszuságok becslésében. Kivételt képeznek természetesen azok az emberek, akik /például foglalkozásuknál fogva/ nagy gyakorlatot szereztek a terület "direkt" becslésében.

Először azzal az esettel foglalkozunk, amikor a területet egy paraméter határozza meg/ négyzet, kör, stb./. Ebben az esetben az információ értékének meghatározása még mindig nem különbözik lényegesen az eddigiektől. A szóbanlevő paramétert, mint valószinüségi változót ξ -vel, a területet, mint valószinüségi változót η -vel jelöljük. A két valószinüségi változó között most a $\eta = c\,\xi^2$ összefüggés áll fenn, ahol c a szóbanlevő terület alakjától függő konstans.

a priori sürüségfüggvényét g/x/-el, posapriori sürüségfüggvényét r/x/-el jelölve, η megfelelő sürüségfüggvényei az alábbi módon irhatók fel:

η a priori sürüségfüggvénye $\dfrac{g\left(\sqrt{\frac{x}{c}}\right)}{2c\sqrt{\frac{x}{x}}}$

poszpriori sürüségfüggvénye $\dfrac{r\left(\sqrt{\frac{x}{c}}\right)}{2c\sqrt{\frac{x}{c}}}$

/ld. PRÉKOPA 84.. old./

A paraméter konkrét értékét h-val, a terület valódi értékét pedig T-vel jelöljük /$T = c \cdot h^2$ /.

Mivel

$$\log \frac{r\ /h/}{g\ /h/} = \log \frac{r\left(\sqrt{\frac{T}{c}}\right)}{g\left(\sqrt{\frac{T}{c}}\right)}$$

, ezért elvileg a η -re vonatkozó információ értéke megegyezik ξ -re kapott információ értékével. Külön feladat lenne e két információ érték gyakorlati kapcsolatának feltárása.

Kissé bonyolultabb a helyzet, ha a területet két paraméter határozza meg. Legyen például ξ egy téglalap szélessége γ pedig a hosszusága, amik a szubjektum számára valószinüségi változók. A területet, mint valószinüségi változót most is η -vel jelöljük. Rögzitett téglalapot feltételezve most is csak szubjektiv sürüségfüggvényekkel dolgozunk. Az egyes valószinüségi változók között most a $\eta = \xi \cdot \zeta$ összefüggés áll fenn.

Vezessük be a következő jelöléseket:

a priori	sürüségfüggvénye	$g_1/x/$,
supriori	sürüségfüggvénye	$r_1/x/$,
a priori	sürüségfüggvénye	$g_2/x/$,
supriori	sürüségfüggvénye	$r_2/c/$.

Ha ξ és ζ függetlenek, akkor η sürüségfüggvényei a következő módon határozhatók meg:

a priori szubjektiv sürüségfüggvénye

$$\int_{-\infty}^{\infty} \frac{1}{y} \, g_1\left(\frac{x}{y}\right) \cdot g_2 \, /y/ \, d\,y \, ,$$

posapriori szubjektiv sürüségfüggvénye

$$\int_{-\infty}^{\infty} \frac{1}{y} \, r_1\left(\frac{x}{y}\right) \cdot r_2 \, /y/ \, d\,y \, .$$

Ezekből a η-ra vonatkozó információ értéke kiszámitható. Az igy számolt információ érték realitásával kapcsolatban ugyanaz a probléma áll fenn, mint az első példánál. E problémát csak kisérleti uton lehet megoldani. A közvetett mérést csak mint lehetőséget emlitettük, az információ értékének meghatározására biztosabb a közvetlen ut.

V. $\mathcal{B}$ -ben tartalmazott $\mathcal{A}$ -ra vonatkozó információ értéke.

Legyen $\mathcal{A}$ és $\mathcal{B}$ két eseményrendszer. Ha a kisérleti személy megtudja $\mathcal{B}$ konkrét kimenetelét, akkor előfordulhat, hogy e közlés az $\mathcal{A}$ -ra vonatkozó bizonytalanságát, szubjektiv entrópiáját csökkenti.

Nézzünk egy egyszerü példát: Két kockával dobunk.

$\mathcal{A}$ eseményrendszer eseményei:

A_1 : a felülre került számok összege 2;

A_2 : a felülre került számok összege 3;

. . .

A_{11} : a felülre került számok összege 12.

$\mathcal{B}$ eseményrendszer eseményei:

B_1 : az első kocka dobásakor 1-es került felülre

B_2 : az első kocka dobásakor 2-es került felülre

. . .

B_6 : az első kocka dobásakor 6-os került felülre.

Ha a kisérletet végrehajtjuk /mindkét kockával dobunk/, akkor a kisérleti személyben $\mathcal{A}$-ra és $\mathcal{B}$ -re vonatkozóan fellép egy szubjektiv valószinüségeloszlás. Ha ezután közöljük a kisérleti személlyel, hogy az első kockával 4-est dobtunk /B_4 következett be/, akkor ez a közlés megváltoztatja az $\mathcal{A}$ -ra eredetileg kialakult szubjektiv valószinüségeloszlást /persze magát az eseményrendszert is, hisz a dobott számok összege most már minimálisan 5 lehet/.

Tehát $\mathcal{B}$ kimenetelének közlése információt ad $\mathcal{A}$-ra. Ez az információ érték attól is függ, hogy $\mathcal{B}$-nek mely eseménye következik be, a $\mathcal{B}$ -ben tartalmazott $\mathcal{A}$-ra vonatkozó információ értékét tehát célszerü egy bizonyos átlaggal mérni.

Természetesen nem minden $\mathcal{A}$ és $\mathcal{B}$ eseményrendszer viszonylatában áll fenn a fenti kapcsolat.

Legyen például $\mathcal{A}$ a következő eseményrendszer:

A_1 : ma esni fog az eső

A_2 : ma nem fog esni az eső

$\mathcal{B}$ megegyezik az előbbivel.

A kisérleti személlyel hiába közöljük $\mathcal{B}$ kimenetelét, ez nem fogja befolyásolni az $\mathcal{A}$-ra kialakult szubjektiv valószinüségeloszlást.

Ebben az esetben tehát $\mathscr{B}$ nem tartalmaz információt $\mathscr{A}$-ra vonatkozóan. Ennek oka, hogy itt nincs kapcsolat a két eseményrendszer között, függetlenek egymástól, az előbbi esetben pedig a két eseményrendszer között bizonyos kapcsolat állt fenn.

Ahhoz, hogy e kérdést matematikailag vizsgáljuk, néhány valószinüségszámitási fogalomra lesz szükségünk. Ezekről röviden szólunk, de még a definiciók pontos leirására sincs lehetőségünk, ezeket az [5.] és [13.] irodalomban találjuk.

Legyen $\mathscr{A} = \{A_i\}$, ennek matematikai eloszlása $\{P(A_i)\}$, szubjektiv eloszlása pedig $\{Q(A_i)\}$, i=1; 2; ...; n.

Ugyanigy $\mathscr{B} = \{B_j\}$, $\{P(B_j)\}$, $\{Q(B_j)\}$, j = 1; 2;...;m.

E két eseményrendszerből egy ujabb eseményrendszert készithetünk, amit $\mathscr{A} \times \mathscr{B}$ -vel jelölünk és a két eseményrendszer direkt szorzatának nevezzük. Definiciója: $\mathscr{A} \times \mathscr{B} = \{A_i \cdot B_j\}$, ahol $A_i \cdot A_j$ jelenti azt az eseményt, hogy A_i és B_j egyszerre következik be. Az $\mathscr{A} \times \mathscr{B}$ n . m eseményből áll.

Például $\mathscr{A}$ és $\mathscr{B}$ legyenek e fejezet első példája szerint definiálva. Ha összesen 9-et dobunk és az elsőn 4-es van felül, akkor azt mondjuk, hogy $\mathscr{A} \times \mathscr{B}$-ből $A_8 \cdot B_4$ esemény következett be.

Az $\mathscr{A}$ és $\mathscr{B}$ eseményrendszert függetlennek nevezzük, ha minden i-re és j-re igaz, hogy $P(A_i \cdot B_j) = P(A_i) \cdot P(B_j)$. E matematikai definició analógiájára két eseményrendszert szubjektive függetleneknek nevezzük, ha minden i-re és j-re fennáll a $Q(A_i \cdot B_j) = Q(A_i) \cdot Q(B_j)$ összefüggés.

Természetesen a matematikai függetlenségből nem következik a szubjektiv függetlenség, de forditva sem, hisz az utóbbi egy egyén elképzelésén, ismeretein alapszik, az előbbi pedig természeti törvényszerüséget fejez ki.

Másik fontos fogalom a feltételes valószinüség fogalma. Legkönnyebben talán a szubjektiv esetben lehet e fogalmat megérteni. $Q(A_i / B_j)$ jelenti A_i esemény szubjektiv valószinüségét, miután a szubjektum megtudta, hogy $\mathcal{B}$ -ből a B_j esemény következett be.

A matematikai feltételes valószinüséget, a $P(A_i / B_j)$ -t is hasonlóan definiáljuk, az előbbi definició teljes átvétele azonban kissé pontatlan definicióhoz vezetne, de a lényegének megértését nem zavarja.

Legkönnyebben a

$$P(A_i / B_j) = \frac{P(A_i \cdot B_j)}{P(B_j)}$$

egyenlőség alapján definiálhatjuk pontosan $P(A_i / A_j)$ -t.

A fenti összefüggés analógiájára fel fogjuk használni a

$$Q(A_i / B_j) = \frac{Q(A_i \cdot B_j)}{Q(B_j)}$$

összefüggést.

Ezzel kapcsolatban azonban felvetődik egy probléma: Különkülön definiáltuk a $Q(A_i / B_j)$, $Q(A_i \cdot B_j)$ és a $Q(B_j)$ fogalmakat. A definiciók alapján mód van ezek felmérésére. Kisérleti adatok hiányában viszont nem tudjuk bizonyitani a fenti összefüggést.

Erre az összefüggésre viszont nagy szükségünk lesz. Az ellentmondás kikerülésére két megoldás kinálkozik: egyik, hogy kisérletileg igazoljuk, esetleg pontositsuk a fenti kapcsolatot, a másik lehetőség pedig, hogy $Q(A_i / B_j)$ -re eleve ezt az összefüggést fogadjuk el definiciónak. Mivel az összefüggés kisérleti igazolására nem vállalkozhatunk, az utóbbi megoldást kell választanunk.

Véleményünk szerint a kisérleti igazolás is pozitiv eredménnyel járna /az esetleges konstans szorzó nem zavarná az alábbi levezetést/, legfeljebb szélsőséges esetekben mutatkozhat lényeges eltérés, aminek feltárása természetesen fontos feladat lenne.

Hasonló fogalmakat folytonos valószinüségi változók esetén is bevezethetünk. Bár erre csak a következő fejezetben lesz szükségünk, a teljesség kedvéért itt szólunk róla.

Legyen ξ és ζ két folytonos valószinüségi változó, $\xi \times \zeta$ jelentse azt az eseményrendszert, melynek elemi eseményét a ξ -nek és ζ -nek egy-egy értéke, mint számpár adja /ezt egy sikvektorral szemléltethetnénk/. /A szokásos jelöléstől az analógia miatt tértünk el./

Ha a két valószinüségi változó független, akkor az együttes eloszlás sürüségfüggvénye a két eloszlás sürüségfüggvényének szorzataként adódik. Itt is beszélhetünk feltételes eloszlásról, amit a feltételes sürüségfüggvény ir le. A számunkra fontos fogalmakról az adott helyen még szólunk, de a tárgyalás egyszerüsége végett igyekszünk kerülni minden uj fogalom bevezetését, főként a már tárgyalt eseményalgebrára igyekszünk támaszkodni.

Ezután rátérhetünk az alcimben szereplő fogalom pontos definiálására:

Első közelitésként I jelentse azt az információt, hogy a $\mathcal{B}$ eseményrendszerből a "B_j esemény következett be". Ekkor az $\mathcal{A}$ eseményrendszer matematikai és szubjektiv eloszlása

$\{P(A_i)\}$ -ről $\{P(A_i/B_j)\}$ -re, illetve $\{Q(A_i)\}$ -ről $\{Q(A_i/B_j)\}$ -re /i=1, 2, ..., n/ változik. A közlés információtartalmát / $\mathcal{A}$ -ra vonatkozóan/ most is a két szubjektiv entrópia különbsége adja:

$$\varnothing_{\mathcal{A}}(I) = \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} - \sum_{i=1}^{n} P(A_i/B_j) \log \frac{1}{Q(A_i/B_j)}$$

Ha a $\mathcal{B}$ kisérletet végrehajtjuk, természetesen nem mindig B_j következik be, tehát nem mindig ekkora információt jelent $\mathcal{B}$ kimenetelének közlése $\mathcal{A}$ -ra. Annak valószinüsége, hogy a fenti információértéket közöljük $P(B_j)$. A fenti információértéket tehát valószinüségi változóként kezelhetjük, ennek várható értékét $\varnothing_{\mathcal{A}}(\mathcal{B})$ -vel jelöljük, és $\mathcal{B}$ -ben tartalmazott $\mathcal{A}$ -ra vonatkozó információnak nevezzük. Köznapibb nyelven ez azt fejezi ki, hogy $\mathcal{B}$ konkrét kimeneteleinek közlésével átlagosan mennyi információt közlünk a szubjektummal.

A fentiek szerint

$$\varnothing_{\mathcal{A}}(\mathcal{B}) = \sum_{j=1}^{m} P(B_j) \left[\sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} - \sum_{i=1}^{n} P(A_i/B_j) \log \frac{1}{Q(A_i/B_j)} \right]$$

Ha felhasználjuk, hogy

$$\sum_{j=1}^{m} P(B_j) = 1, \qquad P(A_i/B_j) \cdot P(B_j) = P(A_i \cdot B_j) ,$$

$$Q(A_i/B_j)\, Q(B_j) = Q(A_i B_j) \text{ és } \sum_{i=1}^{n} P(A_i/B_j) = 1 ,$$

akkor először a

$$= \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} - \sum_{j=1}^{m} \sum_{i=1}^{n} P(A_i \cdot B_j) \log \frac{1}{Q(A_i/B_j)}$$

Alakitsuk még tovább a jobboldali kifejezést:

$$\phi_{\mathcal{A}}(\mathcal{B}) = \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} - \sum_{j=1}^{m} \sum_{i=1}^{n} P(A_i \cdot B_j) \log \frac{Q(B_j)}{Q(A_i \cdot B_j)} =$$

$$= \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} + \sum_{j=1}^{m} \sum_{i=1}^{n} P(A_i/B_j) \cdot P(B_j) \log \frac{1}{Q(B_j)}$$

$$- \sum_{j=1}^{m} \sum_{i=1}^{n} P(A_i \cdot B_j) \log \frac{1}{Q(A_i \cdot B_j)} \quad , \text{ végül}$$

$$\phi_{\mathcal{A}}(\mathcal{B}) = \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} + \sum_{j=1}^{m} P(B_j) \log \frac{1}{Q(B_j)} -$$

$$- \sum_{j=1}^{m} \sum_{i=1}^{n} P(A_i \cdot B_j) \log \frac{1}{Q(A_i \cdot B_j)} \quad .$$

Mivel

$$\phi_{\mathcal{A}}(\mathcal{A}) = \sum_{i=1}^{n} P(A_i) \log \frac{1}{Q(A_i)} \quad , \text{ ezért}$$

$$\phi_{\mathcal{A}}(\mathcal{B}) = \phi_{\mathcal{A}}(\mathcal{A}) + \phi_{\mathcal{B}}(\mathcal{B}) - \phi_{\mathcal{A} \times \mathcal{B}}(\mathcal{A} \times \mathcal{B})$$

Mivel A_i és B_j szerepe felcserélhető, ezért ebből látható, hogy

$$\phi_{\mathcal{A}}(\mathcal{B}) = \phi_{\mathcal{B}}(\mathcal{A})$$

Az is könnyen belátható, hogy ha $\mathcal{A}$ és $\mathcal{B}$ szubjektive függetlenek /a matematikai függőség most lényegtelen/, akkor $\phi_{\mathcal{A}}(\mathcal{B}) = 0$.

Ez azt jelenti, hogy ha a szubjektum nem lát kapcsolatot a két eseményrendszer között, akkor az egyik eseményrendszer kimenetelének közlése nem ad információt a másik eseményrendszerre vonatkozóan.

A fentiekhez hasonló módon járhatnánk el folytonos valószinüségi változó esetében is, ezt azonban nem tesszük meg, mert csak a továbbiakban szükséges alapfogalmak bevezetése, és két eseményrendszer kapcsolatának bemutatása volt a célunk.

E fejezet befejezéseként, de az egész dolgozatra vonatkoztatva ki szeretnénk hangsulyozni, hogy e matematikai formulák pszichológiai és pedagógiai folyamatokat, jelenségeket irnak le, illetve jellemeznek. E formulákhoz hasonlókat találhat az olvasó információfelvétellel foglalkozó matematikai könyvekben, de lényeges az a különbség, hogy mi a szubjektiv valószinüségeket is figyelembe vesszük, igy itt nem matematikai, hanem éppen pszichológiai fogalmak szerepelnek. Mindezt azért hangsulyozzuk ki, mert nem mindig tudunk kitérni a pszichológiai vonatkozásokra, pedagógiai alkalmazásokra. Legtöbbször igen egyszerü /a fenti tudományokban érdektelen/ példákat hozunk, de látni kell, hogy e módszer bonyolultabb pedagógiai folyamat vizsgálatára is alkalmas, aminek elemzése természetesen sokkal hosszadalmasabb.

Ezek illusztrálására felvetünk /ezt is csak vázlatosan/ egy pedagógiai folyamatot, ami éppen e fejezethez illeszkedik.

Közismert az a pszichológiai jelenség, hogy sokkal tartósabb az az ismeret, amit egy szubjektum magától "fedez" fel, mint amit egyszerüen közölnek vele. E tényt a pedagógusok következetesen felhasználják az oktatás folyamatában.

Ha a tanuló egy kérdésre nem tud határozott választ adni, hanem a $\mathcal{A} = \{A_i\}$ válaszokat tartja lehetségesnek, akkor információt kell vele közölni, hogy megtudja, valóban melyik válasz a helyes /az eddigi szóhasználattal, melyik esemény matematikai valószinüsége 1/. Az információ közlése az alábbi módokon is történhet:

1. Közöljük a helyes választ. /Ez pedagógiailag ritkán tekinthető helyes módszernek./
2. $\mathcal{A}$ mellé olyan $\mathcal{B}$, $\mathcal{C}$, ... eseményrendszereket veszünk, melyek szoros /függőségi/ kapcsolatban állnak $\mathcal{A}$-val. Abból, hogy a felvett eseményrendszerekből mely események következnek be, következtet a szubjektum az eredeti kérdés helyes válaszára. /"Rávezető" kérdések feltevéséről van szó/.

Látható, hogy a második eset a most tárgyalt információfelvétellel irható le.

VI. Két eseményrendszer direkt szorzatára vonatkozó információ értéke

Az eddigiakben csak egyetlen $\mathcal{A}$ eseményrendszerre vonatkozó információfelvételről beszéltünk, tehát az információ értékének mérése még nagyon speciális esetre volt értelmezve.

Egy közlés, a szubjektum egy észlelése következtében általában nem csak egy eseményrendszer szubjektiv eloszlása változik meg, hanem többé is.

Az adott közlés információértékét tehát nem mérhetjük az-által, hogy csak egy eseményrendszer szubjektiv valószinüség-eloszlásának változását figyeljük. Most egy lépéssel tovább-megyünk, egyszerre két eseményrendszer, illetve ezek szubjek-tiv eloszlásának változását vesszük számitásba. Ezzel már meg fogjuk mutatni az utat több eseményrendszerre vonatkozó in-formációfelvétel méréséhez is.

Tekintsünk két eseményrendszert, ezeket továbbra is $\mathcal{A} = \{A_i\}$ -vel és $\mathcal{B} = \{B_j\}$ -vel, ezek direkt szorzatát $\mathcal{A} \times \mathcal{B} = \{A_i \cdot B_j\}$ -vel jelöljük, ahol $i = 1, 2, \ldots, n$ és $j = 1, 2, \ldots m$.

Tegyük fel, hogy a szubjektumot olyan hatás éri, hogy mindkét eseményrendszer szubjektiv valószinüségeloszlása megváltozik, a szubjektummal I információt közlünk. Határoz-zuk meg ennek értékét e két eseményrendszerre vonatkozóan.

Nyilvánvalóan helytelenül járnánk el, ha a két esemény-rendszerre külön-külön felmért információértékek összegeként definiálnánk I értékét e két eseményrendszerre vonatkozóan.

Ezt egy példával igyekszünk szemléltetni:

Egy kockával dobunk. A két eseményrendszer legyen a követ-kező:

$\mathcal{A}$: A_1 : 1 -est; A_2 : 2-est; ...
A_6 : 6-ost dobunk.

$\mathcal{B}$: B_1 : 1 -est vagy 2-est;
B_2 : 3-ast vagy 4-est;
B_3 : 5-öst vagy 6-ost dobunk.

I : 3 -ast dobtunk.

Ha feltételezzük, hogy a matematikai és a szubjektiv valószinüségeloszlás megegyezik, akkor

$\varnothing_{\mathcal{A}}(I) = \log 6$; $\varnothing_{\mathcal{B}}(I) = \log 3$, a kettő összege pedig log 18. Ez azt jelentené, hogy a fenti közlés a két eseményrendszerre vonatkozóan log 18 bit információt hordoz. Ekkora információt viszont akkor közlünk, ha 18 egyenlő valószinüségü esemény közül mondjuk meg, hogy melyik következett be. Itt viszont tulajdonképpen elegendő hat eseményt / $\mathcal{A}$ eseményeit/ figyelnünk, mert ha A_1 bekövetkeztét közöljük a szubjektummal, akkor ezzel B_1 bekövetkeztét is közöltük, tehát $\mathcal{A}$ közlésével $\mathcal{B}$ -t is közöljük. Logikus lenne tehát, ha a fenti információnak nem log 18, hanem csak log 6 információ érték felelne meg.

Induljunk ki most is a kezdeti bizonytalanság, illetve a szubjektiv entrópia meghatározásából.

Mivel egyszerre két eseményrendszert figyelünk, ezért tulajdonképpen azt kell szem előtt tartani, hogy melyik A_i, B_j páros következik be. A meglepetés értékét most a

$$\log \frac{1}{Q(A_i \cdot B_j)} \quad \text{-vel definiáljuk,}$$

ennek várható értéke

$$\sum_{i=1}^{n} \sum_{j=1}^{m} P(A_i \cdot B_j) \log \frac{1}{Q(A_i\, B_j)} \text{ , ami nem más, mint}$$

az $\mathcal{A} \times \mathcal{B}$ eseményrendszer szubjektiv entrópiája.

Ha a szubjektummal I információt közlünk, akkor az eseményrendszer szubjektiv eloszlása $R(A_i \cdot B_j)$ -re változik, az uj szubjektiv entrópia tehát

$$\sum_{i=1}^{n}\sum_{j=1}^{m} P(A_i\cdot B_j)\log\frac{1}{\underset{\substack{\uparrow\\ R}}{Q}(A_i\cdot B_j)}\qquad \text{lesz.}$$

E két szubjektiv entrópia különbsége adja az $\mathcal{A}\times\mathcal{B}$ -re vonatkozó információ értékét:

$$\Phi_{\mathcal{A}\times\mathcal{B}}(I)=\sum_{i=1}^{n}\sum_{j=1}^{m} P(A_i\cdot B_j)\log\frac{R(A_i\cdot B_j)}{Q(A_i\cdot B_j)}\;.$$

Alakitsuk tovább ezt a kifejezést:

$$\Phi_{\mathcal{A}\times\mathcal{B}}(I)=\sum_{i=1}^{n}\sum_{j=1}^{m} P(B_j/A_i)\,P(A_i)\log\frac{R(B_j/A_i)\cdot R(A_i)}{Q(B_j/A_i)\cdot Q(A_i)}=$$

$$=\sum_{i=1}^{n} P(A_i)\sum_{j=1}^{n} P(B_j/A_i)\left[\log\frac{R(B_j/A_i)}{Q(B_j/A_i)}+\log\frac{R(A_i)}{Q(A_i)}\right],$$

amiből $\sum_{j=1}^{m} P(B_j/A_i)=1$ felhasználásával a

$$\Phi_{\mathcal{A}\times\mathcal{B}}(I)=\sum_{i=1}^{n} P(A_i)\log\frac{R(A_i)}{Q(A_i)}+$$

$$+\sum_{i=1}^{n} P(A_i)\cdot\sum_{j=1}^{m} P(B_j/A_i)\cdot\log\frac{R(B_j/A_i)}{Q(B_j/A_i)}$$

eredményhez jutunk.

Az első tag $\Phi_{\mathcal{A}}(I)$ -t adja. *A második tagban szereplő*

$$\sum_{j=1}^{m} P(B_j/A_i)\log\frac{R(B_j/A_i)}{Q(B_j/A_i)}$$ mennyiség azt fejezi ki,

hogy I információ mekkora értéket képvisel $\mathcal{B}$ -re, ha már tudjuk, hogy A_i esemény következett be $\mathcal{A}$-ból. Ez az összeg tehát A_i függvénye, annak valószinüsége, hogy a fenti összeggel kell számolnunk $P(A_i)$, a második tag tehát ezen valószinüségi változó várható értéke.

Szavakban elmondva a második tag azt fejezi ki, hogy várhatóan /vagy átlagosan/ I mekkora értéket képvisel $\mathcal{B}$-re vonatkozóan, ha már $\mathcal{A}$ kimenetelét közöltük a szubjektummal. Ezt a mennyiséget $\phi_{\mathcal{B}}(I/\mathcal{A})$ -val fogjuk jelölni.

Tehát

$$\phi_{\mathcal{A}\times\mathcal{B}}(I) = \phi_{\mathcal{A}}(I) + \phi_{\mathcal{B}}(I/\mathcal{A})$$

Egy I információ értékét két eseményrendszerre vonatkozóan $\phi_{\mathcal{A}\times\mathcal{B}}(I)$-vel mérjük. Azt, hogy e mennyiség eleget tesz bizonyos alapvető követelményeknek ezután fogjuk megmutatni, amivel e definició jogosságát alátámasztjuk.

Mivel $\mathcal{A}$ és $\mathcal{B}$ szerepe felcserélhető, ezért a

$$\phi_{\mathcal{A}\times\mathcal{B}}(I) = \phi_{\mathcal{B}}(I) + \phi_{\mathcal{A}}(I/\mathcal{B})$$ alakot is használhatjuk.

Vizsgáljuk még $\phi_{\mathcal{B}}(I/\mathcal{A})$ értéket részletesebben:

$$\phi_{\mathcal{B}}(I/\mathcal{A}) = \sum_{i=1}^{n} P(A_i) \sum_{j=1}^{m} P(B_j/A_i) \log \frac{R(B_j/A_i)}{Q(B_j/A_i)} =$$

$$= \sum_{j=1}^{m} \sum_{i=1}^{n} P(A_i \cdot B_j) \log \frac{R(B_j/A_i)}{Q(B_j/A_i)} =$$

$$= \sum_{j=1}^{m} \sum_{i=1}^{n} P(A_i/B_j) P(B_j) \log \frac{R(B_j/A_i)}{Q(B_j/Ai)} =$$

$$= \sum_{j=1}^{m} P(B_j) \sum_{i=1}^{n} P(A_i/B_j) \log \frac{R(B_j/A_i)}{Q(B_j/A_i)} .$$

Ha a két eseményrendszer szubjektive független, akkor $R(B_j/A_i) = R(B_j)$, illetve $Q(B_j/A_i) = Q(B_j)$, tehát

$$= \sum_{j=1}^{m} P(B_j) \sum_{i=1}^{n} P(A_i/B_j) \log \frac{R(B_j)}{Q(B_j)} =$$

$$= \sum_{j=1}^{m} P(B_j) \log \frac{R(B_j)}{Q(B_j)} \sum_{i=1}^{n} P(A_i/B_j) .$$

Mivel $\sum_{i=1}^{n} P(A_i/B_j) = 1$ minden j-re, ezért szubjektiv függetlenség esetén

$$\phi_{\mathcal{B}}(\mathcal{I}/\mathcal{A}) = \phi_{\mathcal{B}}(\mathcal{I}) \quad \text{, tehát}$$

$$\phi_{\mathcal{A}\times\mathcal{B}}(\mathcal{I}) = \phi_{\mathcal{A}}(\mathcal{I}) + \phi_{\mathcal{B}}(\mathcal{I})$$

Ha viszont bármely $\mathcal{A}$-beli esemény bekövetkeztének közlése után tudja a szubjektum, hogy $\mathcal{B}$-ből melyik esemény fog bekövetkezni /mint a bevezető példánk esetében/, akkor $R(B_j/A_i)$ és $Q(B_j/A_i)$ egyszerre 0 vagy 1 értéket vesz fel. Ebben az esetben

$\phi_{\mathcal{B}}(I/\mathcal{A}) = 0$, tehát $\mathcal{A}$ közlése után minden információ értéke 0.

Az előbbi példánknál például ha tudjuk, hogy melyik számot dobtuk, akkor azt is tudjuk, hogy melyik kettő közül dobtuk.

Igy az ottani sejtésünk matematikai megfogalmazást nyert, $\phi_{\mathcal{A}\times\mathcal{B}}(I)$ tehát nem log 18 /mivel abban az esetben $\phi_{\mathcal{B}}(I/\mathcal{A})=0$ /hanem log 6, amit vártunk.

Az általános eset tárgyalásakor most is /mint egyetlen $\mathcal{A}$ eseményrendszer esetében/ figyelembe kell venni, hogy I közlése után nemcsak a szubjektiv eloszlás változhat meg, hanem egy uj matematikai eloszlással, uj valószinüségi mezővel kell dolgoznunk.

Ilyen esetben az $\mathcal{A}\times\mathcal{B}$ -re vonatkozó információ értékét a

$$\phi_{\mathcal{A}\times\mathcal{B}}(I)=\sum_{i=1}^{n}\sum_{j=1}^{m} P(A_i.B_j)\log\frac{1}{Q(A_i.B_j)} -$$

$$-\sum_{k=1}^{u}\sum_{l=1}^{v} P'(A'_k\ B'_l)\log\frac{1}{Q'(A'_k\ B'_l)}$$

képlet alapján számolhatjuk, ahol $\{A'_k\}$ és $\{B'_l\}$ az uj eseményrendszereket, P' és Q' pedig az uj valószinüségeket jelölik.

Megfelelő /itt nem részletezett/ átalakitások esetén $\phi_{\mathcal{A}\times\mathcal{B}}(I)$ a következő alakra hozható:

$$\phi_{\mathcal{A}\times\mathcal{B}}(I)=\left\{\sum_{i=1}^{n} P(A_i)\log\frac{1}{Q(A_i)}-\sum_{k=1}^{u} P'(A_k)\log\frac{1}{Q'(Ak)}\right\}+$$

$$+\left\{\sum_{i=1}^{n} P(A_i)\sum_{j=1}^{m} P(B_j/A_i)\log\frac{1}{Q(B_j/A_i)} -\right.$$

$$- \sum_{k=1}^{u} P'(A_k) \sum_{l=1}^{v} P'(B'_l/A'_k) \log \frac{1}{Q'(B'_l/A'_k)} \Bigg\} .$$

Az első zárójelben levő különbség $\Phi_{\mathcal{A}}(\mathcal{J})$ -t adja.

A második zárójelben levő különbség ugyancsak szubjektiv entrópiákból épül fel. Pontosabban szubjektiv entrópiák várható értékeinek különbsége szerepel a második zárójelben. Ennek szavakban való megfogalmazása nehézkes. Röviden ugy mondhatjuk, hogy a különbség első tagja azon szubjektiv entrópiák várható értéke, amelyek A_i közlése után /de még I közlése előtt/ fennmaradnak. A második tag /a kivonandó/ azon szubjektiv entrópiák várható értéke, melyek A'_k események bekövetkeztének közlése után maradnak fenn, miután I-t is közöltük. Ez természetes általánositása a már szerepelt $\Phi_{\mathcal{B}}(\mathcal{J}/\mathcal{A})$ - nak, ezért ezt ugyanigy jelöljük. Tehát most is igaz, hogy

$$\Phi_{\mathcal{A}\times\mathcal{B}}(\mathcal{J}) = \Phi_{\mathcal{A}}(\mathcal{J}) + \Phi_{\mathcal{B}}(\mathcal{J}/\mathcal{A})$$

$\Phi_{\mathcal{A}\times\mathcal{B}}(\mathcal{J})$ legfontosabb /már bemutatott/ tulajdonságai:

1. $\Phi_{\mathcal{A}\times\mathcal{B}}(\mathcal{J}) = \Phi_{\mathcal{A}}(\mathcal{J}) + \Phi_{\mathcal{B}}(\mathcal{J})$, ha $\mathcal{A}$ és $\mathcal{B}$ függetlenek

2. $\Phi_{\mathcal{A}\times\mathcal{B}}(\mathcal{J}) = \Phi_{\mathcal{A}}(\mathcal{J})$, ha $\mathcal{A}$ közlése maga után vonja $\mathcal{B}$ közlését is.

E két tulajdonság indokolja, hogy $\Phi_{\mathcal{A}\times\mathcal{B}}(\mathcal{J})$-t választhatjuk a két eseményrendszerre vonatkozó információ értékének.

Folytonos valószinüségi változók esetében ezeket a gondolatmeneteket kell megismételni, persze a megfelelő átfogalmazással.

ξ és ζ legyen két - sürüségfüggvénnyel rendelkező - valószinüségi változó. Határozzuk meg $\Phi_{\xi\times\zeta}(\eta)$ értékét / $\xi\times\zeta$ értelmezését az V. fejezetben adtuk meg/. Alapdefiníció most a

$$\Phi_{\xi\times\zeta}(\eta) = \lim_{\substack{n\to\infty \\ m\to\infty}} \sum_{i=1}^{n} \sum_{j=1}^{m} P\left(X_i \le \xi \le X_{i+1}\ ;\ Y_j \le \zeta \le Y_{j+1}\right) \cdot$$

$$\cdot \log \frac{R\left(X_i \le \xi \le X_{i+1}\ ;\ Y_j \le \zeta \le Y_{j+1}\right)}{Q\left(X_i \le \xi \le X_{i+1}\ ;\ Y_j \le \zeta \le Y_{j+1}\right)} \quad \text{határérték.}$$

Az $\left(X_i \le \xi \le X_{i+1};\ Y_j \le \zeta \le Y_{j+1}\right)$ jelölés azt fejezi ki, hogy ξ X_i és X_{i+1} között, ugyanakkor ζ Y_j és Y_{j+1} között van, vagyis e két feltétel egyszerre teljesül.

Most a diszkrét esetben leirt átalakitásokat kellene megismételnünk, például

$Q\left(X_i \le \xi \le X_{i+1};\ Y_j \le \zeta \le Y_{j+1}\right)$ helyett

$Q\left(Y_j \le \zeta \le Y_{j+1}\ /\ X_i \le \xi \le X_{i+1}\right) Q\left(X_i \le \xi \le X_{i+1}\right)$ -t irva.

Ezt itt nem részletezzük, csak az átalakitás után kapott értéket közöljük.

$$\phi_{\xi\times\zeta}(\mathcal{J}) = \lim_{n\to\infty} \sum_{i=1}^{n} P\left(X_i \le \xi \le X_{i+1}\right) \log \frac{R\left(X_i \le \xi \le X_{i+1}\right)}{Q\left(X_i \le \xi \le X_{i+1}\right)} +$$

$$+ \lim_{\substack{n\to\infty \\ m\to\infty}} \sum_{i=1}^{n} P\left(X_i \le \xi \le X_{i+1}\right) \sum_{j=1}^{m} P\left(Y_j \le \zeta \le Y_{j+1} / X_i \le \xi \le X_{i+1}\right) \cdot$$

$$\cdot \log \frac{R\left(Y_j \le \zeta \le Y_{j+1} \;/\; X_i \le \xi \le X_{i+1}\right)}{Q\left(Y_j \le \zeta \le Y_{j+1} \;/\; X_i \le \xi \le X_{i+1}\right)}$$

Az alapdefinició, vagy annak átalakitott formája a lim jel elhagyása után már alkalmas $\phi_{\xi\times\zeta}(\mathcal{J})$ közelitő meghatározására, sőt elfőrdulhat, hogy gyakorlatilag csak ezek használhatók. A megfelelő eloszlásfüggvények /a ξ és ζ együttes eloszlásfüggvényeit is ide sorolva/ ismeretében a fenti összeg zárt alakban is felirható.

A zárt alak felirását ugy kezdjük, hogy az összegben szereplő valószinüségértékeket az egyes eloszlásfüggvények integráljaként irjuk fel, majd meghatározzuk az összeg határértékét, miközben $n \to \infty$ és $m \to \infty$.

Ha a szükséges müveleteket elvégezzük, akkor az első tagra a már ismert

$$\int_a^b f/x/ \; \log \frac{r/x/}{g/x/} \, dx \quad \left(= \phi_\xi(\mathcal{J})\right)$$

kifejezést, a másikra pedig az

$$\int_a^b \int_d^c h\,/x;y/\,\log \frac{v/x;y/\,.\,g/x/}{u/x;y/\,.\,r/x/}\,dy\,dx$$

$$= \phi_\eta(\eta/\xi)$$

kifejezést kapjuk, ahol h/x;y/ a matematikai, u/x;y/ az eredeti szubjektiv és v/x;y/ pedig az uj szubjektiv együttes eloszlásfüggvénye ξ -nek és η -nek.

Tehát

$$\phi_{\xi\times\eta}(\eta) = \phi_\xi(\eta) + \phi_\eta(\eta/\xi)$$

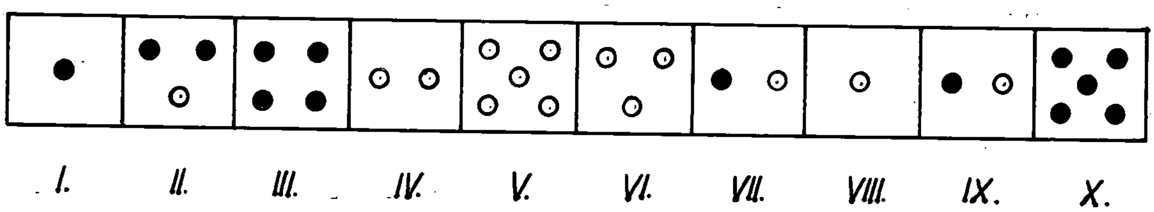
Az előbb kiemelt két alaptulajdonság most is fennáll:

1. Ha ξ és η szubjektive függetlenek, akkor
   u /x;y/ = g/x/ . k/x/ és
   v /x;y/ = r/x/ . l/x/ , ami alapján könnyen belátható, hogy
   $\phi_\eta(\eta/\xi) = \phi_\eta(\eta)$, vagyis
   $$\phi_{\xi\times\eta}(\eta) = \phi_\xi(\eta) + \phi_\eta(\eta)$$
2. Az eredeti definiciót átalakitva juthatnánk könnyen /ugyanugy, mint diszkrét esetben/ olyan alakhoz, amiből kitünne, hogy ha ξ konkrét eredményének közlésével a szubjektum megtudja η kimenetelét is, akkor

$$\phi_\eta(\eta/\xi) = 0 \;.$$

Két eseményrendszerre vonatkozó információ értékének kiszámitására olyan példát mutatunk be, mely egyrészt megmutatja a számolás tényleges elvégzését, másrészt egy érdekes tulajdonságot is szemléltet.

Helyezzünk 10 /római számokkal jelölt/ fiókba 28 golyót, melyek közül 14 fehér és 14 fekete, elosztásuk a 9.ábrán látható.



9.ábra

Tegyük fel, hogy a golyókat 1-től 28-ig sorszámoztuk, de ezt a fiókokba történő elosztásnál nem kell figyelnünk.

Egy urnából /melyben 1-től 28-ig számozott lapok vannak elhelyezve/ huzzuk ki valamely golyó számát.

Két kérdést teszünk fel a kisérleti személynek:

a./ Mit gondol, milyen szinü golyó számát huztuk?

b./ Melyik fiókban van a kihuzott számu golyó?

Mindkét kérdés egy-egy eseményrendszert határoz meg. Az egyes események matematikai valószinüségeit a kedvező és az összes esetek számának hányadosai adják.

$\mathcal{A}$ eseményrendszer:

A_1 /fekete/ A_2 /fehér/

$P/A_1/ = 0,5$ $P/A_2/ = 0,5$

$\mathcal{B}$ eseményrendszer:

B_i jelenti, hogy a kihuzott számu golyó az i-edik fiókban van.

A $\mathcal{B}$ eseményrendszer matematikai valószinüségeloszlását az alábbi táblázat tartalmazza:

i	1	2	3	4	5
$P/A_i/$	0,036	0,107	0,143	0,071	0,179

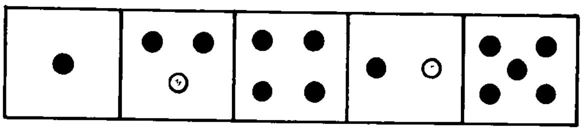
i	6	7	8	9	10
$P/A_i/$	0,107	0,071	0,036	0,071	0,179

Tegyük fel, hogy a szubjektiv valószinüségeloszlás megegyezik a matematikaival. Ez a feltétel nem befolyásolja a példa célját, hisz a számolás más esetben is az alábbiak szerint történne, csak a log jel előtt levő szorzó nem egyezne meg a log jel alatt levő tört nevezőjével.

Tegyük fel továbbá, hogy a szám kihuzása után a kisérleti személy a következő információt kapja:

"A kihuzott szám által meghatározott golyó nincs a IV., V., VI., VII. és VIII. fiókban."

Ekkor két uj valószinüségi mezőt kell figyelembe vennünk /az elsőnél ugyan az események nem változtak, de a matematikai valószinüségek igen, igy ezt is uj valószinüségi mezőként kell kezelnünk/. A megmaradt lehetőségeket a 10. ábra szemlélteti:



10. ábra

$\mathcal{A}'$ eseményrendszer:

A' /fekete/ — A'_2 /fehér/

$P'/A'_1/ = 0,867$ — $P'/A'_2/ = 0,133$

/Mivel 2 fehér golyó maradt és összesen 15 golyó jöhet számitásba, ezért

$$P'/A'_2/ = \frac{2}{15} = 0,133 \ ./$$

$\mathcal{B}'$ eseményrendszer:

$P'/B'_1/ = 0,067$; $P'/B'_2/ = 0,2$; $P'/B'_3/ = 0,267$

$P'/B'_4/ = 0,133$; $P'/B'_5/ = 0,333$.

Most is tegyük fel, hogy a matematikai és a szubjektiv valószinüségeloszlás megegyezik.

I értéke $\mathcal{A}$ -ra és $\mathcal{B}$ -re a már leirt módon számolható, ezekre az alábbi eredményeket kapjuk:

$\Phi_{\mathcal{A}}(I) = 0,435$; $\Phi_{\mathcal{B}}(I) = 0,989$.

Feladatunk I-nek az $\mathcal{A} \times \mathcal{B}$ -re vonatkozó értékének meghatározása.

Ezt kétféleképpen is kiszámithatjuk:

1./ Meghatározzuk $P(A_i \cdot B_j) = Q(A_i \cdot B_j)$ és $P'(A'_k \cdot B'_l) = Q'(A'_k \cdot B'_l)$ értékeket, majd az alap-definició alapján számitjuk $\Phi_{\mathcal{A}\times\mathcal{B}}(I)$ értékét.

2./ Kiszámitjuk a $P(A_i/B_j) = Q(A_i/B_j) = A_i/B_j$ és $P'(A'_k // B'_l) = Q'(A'_k / B'_l)$ értékeket, amikből $\Phi_{\mathcal{A}}(I/\mathcal{B})$ nyerhető, és ezt a már kiszámitott $\Phi_{\mathcal{B}}(I)$ -hez adva megint $\Phi_{\mathcal{A}\times\mathcal{B}}(I)$ -t kapjuk.

Mindkét utat bemutatjuk.

1./ $P(A_1 \cdot B_1) = 0{,}036$, mivel $P(B_1) = 0{,}036$ és B_1 bekövetkezése maga után vonja A_1 bekövetkezését, mert az első fiókban csak fekete golyó van. Igy együttes bekövetkezésüknek valószinüsége megegyezik B_1 valószinüségével.

$P(A_1 \cdot B_2) = 0{,}071$, mivel $P(B_2) = 0{,}107$, de a második fiókban két fekete és egy fehér golyó van, ezért a két fajta golyó számának arányával, $\frac{2}{3}$ -al kell szorozni a 0,107-et, hogy $A_1 \cdot B_2$ valószinüségét kapjuk.

$P(A_1 \cdot B_4) = 0$, mert a negyedik fiókban nincs fekete golyó, e két esemény tehát együtt nem következhet be.

Hasonló módon kapjuk a többi valószinüségértékeket is.

$\mathcal{A} \times \mathcal{B}$ esetében.

$P(A_1 \cdot B_1) = 0{,}036$	$P(A_2 \cdot B_1) = 0$
$P(A_1 \cdot B_2) = 0{,}071$	$P(A_2 \cdot B_2) = 0{,}036$
$P(A_1 \cdot B_3) = 0{,}143$	$P(A_2 \cdot B_3) = 0$
$P(A_1 \cdot B_4) = 0$	$P(A_2 \cdot B_4) = 0{,}071$
$P(A_1 \cdot B_5) = 0$	$P(A_2 \cdot B_5) = 0{,}179$
$P(A_1 \cdot B_6) = 0$	$P(A_2 \cdot B_6) = 0{,}107$
$P(A_1 \cdot B_7) = 0{,}036$	$P(A_2 \cdot B_7) = 0{,}036$
$P(A_1 \cdot B_8) = 0$	$P(A_2 \cdot B_8) = 0{,}036$
$P(A_1 \cdot B_9) = 0{,}036$	$P(A_2 \cdot B_9) = 0{,}036$
$P(A_1 \cdot B_{10}) = 0{,}179$	$P(A_2 \cdot B_{10}) = 0$

$\mathcal{A}' \times \mathcal{B}'$ esetében:

$P'(A_1' \cdot B_1') = 0{,}067$	$P'(A_2' \cdot B_1') = 0$
$P'(A_1' \cdot B_2') = 0{,}133$	$P'(A_2' \cdot B_2') = 0{,}067$
$P'(A_1' \cdot B_3') = 0{,}267$	$P'(A_2' \cdot B_3') = 0$
$P'(A_1' \cdot B_4') = 0{,}067$	$P'(A_2' \cdot B_4') = 0{,}067$
$P'(A_1' \cdot B_5') = 0{,}333$	$P'(A_2' \cdot B_5') = 0$

$$\Phi_{\mathcal{A}\times\mathcal{B}}(\mathcal{I}) = \left[0{,}036 \cdot \log \frac{1}{0{,}036} + \dots + 0 \cdot \log \frac{1}{0}\right] -$$

$$- \left[0{,}067 \cdot \log \frac{1}{0{,}067} + \dots + 0 \cdot \log \frac{1}{0}\right]$$

A részletszámitásoket elvégezve

$\Phi_{\mathcal{A}\times\mathcal{B}}(\mathcal{I}) = 0{,}919$ eredményhez jutunk.

Az eredmény érdekessége abban áll, hogy ez kisebb $\varnothing_A(1)$-nél is és $\varnothing_B(1)$ -nél is. Ennek magyarázatát a másik számitási mód bemutatása utján adjuk meg.

2./ Határozzuk meg $\varnothing_A(1/B)$ értékét.

$P(A_1/B_1) = 1$, mert az első fiókban csak fekete golyó van, tehát B_1 bekövetkezésekor biztosan A_1 is bekövetkezik.

$P(A_1/B_2) = 0{,}666$, mert a második fiók $\frac{2}{3}$ / = 0,666/ - része fekete, tehát ha tudjuk, hogy ebből huztunk, akkor 0,666 annak a valószinüsége, hogy feketét huztunk.

$P(A_1/B_4) = 0$, mivel a negyedik fiókban nincs fekete golyó.

Hasonló gondolkodással kapjuk a többi feltételes való-szinüségértékeket is.

$P(A_1 / B_1) = 1$	$P(A_2 / B_1) = 0$
$P(A_1 / B_2) = 0{,}666$	$P(A_2 / B_2) = 0{,}333$
$P(A_1 / B_3) = 1$	$P(A_2 / B_3) = 0$
$P(A_1 / B_4) = 0$	$P(A_2 / B_4) = 1$
$P(A_1 / B_5) = 0$	$P(A_2 / B_5) = 1$
$P(A_1 / B_6) = 0$	$P(A_2 / B_6) = 1$
$P(A_1 / B_7) = 0{,}5$	$P(A_2 / B_7) = 0{,}5$
$P(A_1 / B_8) = 0$	$P(A_2 / B_8) = 1$
$P(A_1 / B_9) = 0{,}5$	$P(A_2 / B_9) = 0{,}5$
$P(A_1 / B_{10}) = 1$	$P(A_2 / B_{10}) = 0$

$\mathcal{A}'\times\mathcal{B}'$-re vonatkozóan:

$P'(A'_1 / B'_1) = 1$	$P'(A'_2 / B'_1) = 0$
$P'(A'_1 / B'_2) = 0,666$	$P'(A'_2 / B'_2) = 0,333$
$P'(A'_1 / B'_3) = 1$	$P'(A'_2 / B'_3) = 0$
$P'(A'_1 / B'_4) = 0,5$	$P'(A'_2 / B'_4) = 0,5$
$P'(A'_1 / B'_5) = 1$	$P'(A'_2 / B'_5) = 0$

A

$$\Phi_{\mathcal{A}}(\mathcal{I}/\mathcal{B}) = \sum_{j=1}^{m} P(B_j) \sum_{i=1}^{n} P(A_i/B_j) \cdot \log \frac{1}{Q(A_i/B_j)} -$$

$$- \sum_{l=1}^{v} P'(B'_l) \sum_{k=1}^{u} P'(A'_k / B'_l) \cdot \log \frac{1}{Q'(A'_k/B'_l)}$$

képlet alapján

$\Phi_{\mathcal{A}}(\mathcal{I}/\mathcal{B}) = - 0,07$ eredményt kapunk.

Tehát

$\Phi_{\mathcal{A}\times\mathcal{B}}(\mathcal{I}) = 0,989 - 0,07 = 0,919$, ami megegyezik az előbbi eredménnyel.

Mi a magyarázata annak, hogy $\mathcal{B}$ közlése után I már negativ információt szolgáltat? Az átlag negativ értéke érthetőbb lesz, ha látjuk, hogy egyes esetekben I már negativ információ értékeket szolgáltat. Ilyen példát viszont könnyü mutatni.

Tegyük fel, hogy B_5 következett be $\mathcal{B}$-ből. Ennek közlése után $\mathcal{A}$-ra 0 értékü szubjektiv entrópia marad fent.

Ha viszont ezután közöljük I-t, akkor $\mathcal{A}$ szubjektiv entrópiája 0,565 lesz. A két entrópia különbsége adja a felvett információ értékét:

$$\Phi_{\mathcal{A}}(I / B_5) = 0 - 0,565 = -0,565$$

Ennek tartalma az, hogy B_5 igaz volta esetén I már félrevezetést jelent.

Más irányu megközelitésben ugy fogalmazhatunk, hogy az eredeti $\mathcal{A} \times \mathcal{B}$ eseményrendszer esetében $\mathcal{B}$ egy konkrét kimenetelének közlése átlagosan kisebb szubjektiv entrópiát hagy, mint az uj $\mathcal{A}' \times \mathcal{B}'$ esetében a $\mathcal{B}'$ kimenetelének közlése /ld. .115.. oldal/.

Természetesen ez nem általános jelenség./Különben $\Phi_{\mathcal{B}}(I/\mathcal{A})$ itt is pozitiv érték./ Az $\mathcal{A}, \mathcal{B}$ eseményrendszer-párra adott információérték definiciójában azt a fontos tulajdonságot kell észrevenni, hogy ez számitásba veszi I-nek $\mathcal{A}$ -ra vonatkozó értékét, és ehhez adódik még az az információérték, ami $\mathcal{B}$ -re vonatkozik, de az $\mathcal{A}$ -ra számitott értéknél még nem vettünk figyelembe.

Még egyszer hangsulyozzuk e definició - már leirt - két tulajdonságának fontosságát.

VII. Több eseményrendszer direkt szorzatára vonatkozó információ értéke

Az előbbi fejezetben bevezetett fogalmat fogjuk általánositani.

Vegyük fel az $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_s$ eseményrendszereket. Egy I információ értékét e eseményhalmazra vonatkozóan most a $\phi_{\mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_s}(\mathcal{I})$ -vel definiáljuk.

Mivel az egyes események között értelmezett szorzás kommutativ müvelet / A . B = B . A /, ezért

$$\phi_{\mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_s}(\mathcal{I}) = \phi_{\mathcal{A}_s \times \mathcal{A}_{s-1} \times \dots \times \mathcal{A}_2 \times \mathcal{A}_1}(\mathcal{I}).$$

Tekintsük $\mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_s$ -et egy eseményrendszernek, és alkalmazzuk a már levezetett

$\phi_{\mathcal{A} \times \mathcal{B}}(\mathcal{I}) = \phi_{\mathcal{B}}(\mathcal{I}) + \phi_{\mathcal{A}}(\mathcal{I}/\mathcal{B})$ összefüggést:

$$\phi_{\mathcal{A}_s \times (\mathcal{A}_{s-1} \times \dots \times \mathcal{A}_2 \times \mathcal{A}_1)}(\mathcal{I}) = \phi_{\mathcal{A}_{s-1} \times \dots \times \mathcal{A}_2 \times \mathcal{A}_1}(\mathcal{I}) + \phi_{\mathcal{A}_s}(\mathcal{I}/\mathcal{A}_{s-1} \times \dots \times \mathcal{A}_2 \times \mathcal{A}_1).$$

A jobboldalon levő első tagra ugyanezt a felbontást elvégezve, majd az eljárást folytatva kapjuk, hogy

$$\phi_{\mathcal{A}_s \times \mathcal{A}_{s-1} \times \dots \mathcal{A}_2 \times \mathcal{A}_1}(\mathcal{I}) = \phi_{\mathcal{A}_1}(\mathcal{I}) + \phi_{\mathcal{A}_2}(\mathcal{I}/\mathcal{A}_1) + \phi_{\mathcal{A}_3}(\mathcal{I}/\mathcal{A}_1 \times \mathcal{A}_2) + \dots + \phi_{\mathcal{A}_s}(\mathcal{I}/\mathcal{A}_{s-1} \times \dots \times \mathcal{A}_1),$$

illetve

$$\phi_{\mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_s}(\mathcal{I}) = \phi_{\mathcal{A}_1}(\mathcal{I}) + \phi_{\mathcal{A}_2}(\mathcal{I}/\mathcal{A}_1) + \phi_{\mathcal{A}_3}(\mathcal{I}/\mathcal{A}_1 \times \mathcal{A}_2) + \dots + \phi_{\mathcal{A}_s}(\mathcal{I}/\mathcal{A}_1 \times \dots \times \mathcal{A}_{s-1}).$$

E felbontásnak több eseményrendszer esetében már nagyobb jelentősége van, mint ha csak két eseményrendszerre vonatkozó információ értékét számitjuk. Az előbbi fejezet példájánál ugyanis még nem volt lényeges különbség /a számolás nehézségét tekintve/ az eredeti definició és a felbontás alapján történő számolás között.

Több eseményrendszer esetében a felbontás előnye akkor mutatkozik, ha az egyes események kimenetelének közlése után I már nem tartalmaz lényeges információt, tehát az összeg tagjainak abszolutértéke egyre csökken, esetleg bizonyos tagtól kezdve elhanyagolhatók lesznek. Ilyen eset a pedagógiai gyakorlatban sürün fordul elő, példánk is ilyen természetü lesz.

A felbontás másik, elvi szempontból fontosabb jelentősége, hogy igy /megszámlálhatóan/ végtelen sok eseményrendszerre vonatkozó információértéket is definiálhatunk. Ebben az esetben a következő számsor definiálja az információ értékét:

$$\varnothing_{\mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_s \times \dots}(\mathcal{I}) = \varnothing_{\mathcal{A}_1}(\mathcal{I}) + \varnothing_{\mathcal{A}_2}(\mathcal{I}/\mathcal{A}_1) +$$

$$+ \varnothing_{\mathcal{A}_3}(\mathcal{I}/\mathcal{A}_1 \times \mathcal{A}_2) + \dots + \varnothing_{\mathcal{A}_s}(\mathcal{I}/\mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_{s-1}) + \dots$$

Ez az összeg csak akkor ad véges eredményt, ha a számsor konvergens. A határérték pontos meghatározásához az kell, hogy az egymás után következő tagok matematikai törvényszerüség szerint kövessék egymást. Ez a legritkább esetben áll fenn, a törvényszerüség megállapitása ekkor sem könnyü feladat. Ellenkező esetben, amikor matematikai törvényszerüséget nem fedezünk fel, az összeg határértékét közelitő módszerrel határozhatjuk meg.

A gyakorlati számitásnál ki kell választani azokat az eseményrendszereket, melyekre I lényeges hatással van /amelyeknek szubjektiv valószinüségeloszlása a többitől függetlenül lényegesen változik/ és csak ezekre kell I értékét kiszámitani. Ez az érték közelitőleg az összes eseményrendszerre vonatkozó információértéket adja. Ezt nem mindig tehetjük meg, előfordulhat, hogy végtelen sok eseményrendszerre van I lényeges hatással. Erre is mutatunk példát.

Ez a definició kontinuum számosságu /sorba nem szedhető/ eseményrendszer halmaz esetében nem használható. Ilyen esetben véges, vagy megszámlálhatóan végtelen eseményrendszer halmazt kell /ha lehet/ ugy kiválasztani, hogy a kiválasztottakra kiszámitott információérték jól közelitse az összes eseményrendszerre vonatkozó információértéket.

Meg kell még jegyezni, hogy folytonos esetben az előző fejezetben bemutatott módon juthatnánk el az általános definicióhoz. Már ott is problémát okozott az együttes eloszlás sürüségfüggvényének meghatározása, ami itt fokozottan jelentkezne. Matematikailag természetesen járható az az ut is, ha erre az elméleti felépitésnél szükség van, gyakorlati szempontból azonban nincs nagy jelentősége, ezért itt már nem irjuk le. Gyakorlati számitásoknál a folytonos esetet célszerü visszavezetni a diszkrét esetre ugy, hogy adott intervallumokat, illetve /mivel most többdimenziós esetet tárgyalunk/ tartományokat kell elemi eseményeknek tekintenünk.

A fenti képlet megvilágitására, a gyakorlati számolás módjának megmutatására ismét egy konkrét példát mutatunk be.

E példa arra is utal, hogy módszerünket nem csak speciális pszichológiai kisérleteknél, hanem a pedagógiai gyakorlathoz igen közel eső területeken is alkalmazhatjuk.

Tegyük fel, hogy matematika órán a négyzetgyök fogalmát kell tisztázni olyan diákok számára, akik még erről nem hallottak.

A tanár például először a $\sqrt{16}$; $\sqrt{49}$ és a $\sqrt{121}$ értékét akarja a diákokkal tudatni. Legegyszerübb eljárás, ha a tanár megmondja a négyzetgyök definicióját és ez alapján válaszolnak a feltett kérdésekre a diákok. E módszer jó képességü osztály esetében nem okoz problémát, elfogadható. Gyengébb osztálynál azonban célszerübb a probléma felvetése után fokozatosan, a helytelen válaszokat kiszürve ugy irányitani a tanulók gondolkodását, hogy ők maguk /ha először pontatlanul is/ fogalmazzák meg a definiciót.

Tegyük fel, hogy egy tanár a következőképpen jár el: felirja, hogy $\sqrt{25} = 5$, majd megkérdezi az osztályt, hogy milyen értéket adnának a $\sqrt{16}$; $\sqrt{49}$ és $\sqrt{121}$ -re.

Válasszunk ki egy tanulót, akiben a következő eseményrendszerek léptek fel:

$\mathcal{A}$	A_1	A_2	A_3	A_4
	$\sqrt{16} = 6$	$\sqrt{16} = 4$	$\sqrt{16} = \frac{16}{5}$	Egyik sem
$P(A_i)$	0	1	0	0
$Q(A_i)$	0,3	0,4	0,1	0,2

$\mathcal{B}$	B_1	B_2	B_3	B_4
	$\sqrt{49} = 9$	$\sqrt{49} = 7$	$\sqrt{49} = \frac{49}{5}$	Egyik sem
$P(B_j)$	0	1	0	0
$Q(B_j)$	0,3	0,4	0,1	0,2

$\mathcal{C}$	C_1	C_2	C_3	C_4	C_5
	$\sqrt{121}=21$	$\sqrt{121}=1$	$\sqrt{121}=11$	$\sqrt{121}=\frac{121}{5}$	Egyik sem
$P(C_r)$	0	0	1	0	0
$Q(C_r)$	0,2	0,2	0,2	0,1	0,1

Ugy gondoljuk, hogy ezek a szubjektiv eloszlások valóban felléphetnek, hisz például a $\sqrt{25} = 5$ -ből még egyaránt következhet a $\sqrt{49} = 9$ és a $\sqrt{49} = 7$ is.

Hogy a szubjektiv eloszlás közelebb kerüljön a matematikaihoz, a tanárnak ujabb információt kell közölnie a tanulókkal. Az információ megválasztása már az oktatás folyamatában didaktikai kérdést jelent. Abszolutértékben helyes, didaktikailag mégis helytelen információ lenne például a $\sqrt{625} = 25$ közlése. Ekkor ugyanis éppen a gyengébb tanulók számára negativ információt közölnénk, hisz ők hajlamosak a könnyebb megfogalmazásu definició elfogadására, jelen esetben azt a következtetést vonnák le, hogy "a $\sqrt{\ }$ -jel azt jelenti, hogy az alatta levő szám első jegyét el kell hagyni". E következtetés nálunk negativ információértékként jelentkezne.

Bár e példa szépen mutatja, hogy az információfelvétel mérésével az oktatás módszerének helyességére is következtethetünk, mégsem hisszük, hogy a legjobb információ kiválasztását ilyen számitásoknak kell megelőznie. Biztosak vagyunk azonban abban, hogy az elméleti ismeretek elősegitik a gyakorlati munkát. Jelen esetben tudnia kell a tanárnak, hogy a rosszul választott információ /még ha helyes is/ negativ értékü lehet.

Tegyük fel, hogy az I információt a $\sqrt{64} = 8$ közlése jelenti, és a következő valószinüségi mezők alakulnak ki hatására:

$\mathcal{A}'$	A_1'	A_2'
	$\sqrt{16} = 4$	$\sqrt{16} \neq 4$
$P'(A_i')$	1	0
$Q'(A_i')$	0,6	0,4

$\mathcal{B}'$	B_1'	B_2'
	$\sqrt{49} = 9$	$\sqrt{49} \neq 9$
$P'(B_j')$	1	0
$Q'(B_j')$	0,6	0,4

$\mathcal{C}'$	C_1'	C_2'
	$\sqrt{121} = 11$	$\sqrt{121} \neq 11$
$P'(C_r)$	1	0
$Q'(C_r)$	0,6	0,4

/Ezek a szubjektiv eloszlások azt fejezik ki, hogy a tanuló még mindig bizonytalan. Bár a két példa esetében az eredmény négyzete a $\sqrt{}$-jel alatti szám, de előfordulhat, hogy ez nem mindig van igy./

A fenti valószinüségértékek még nem elegendők az információ értékének kiszámitásához, a megfelelő feltételes matematikai és szubjektiv valószinüségértékeket is ismerni kell.

Mi az alábbi adatokkal számolunk:

$P(B_2 / A_2) = 1$ $\qquad$ $P'(B_1' / A_1') = 1$

$Q(B_2 / A_2) = 0,6$ $\qquad$ $Q'(B_1' / A_1') = 0,8$

$P(C_2 / A_2 \cdot C_3) = 1$ $\qquad$ $P'(C_1' / A_1' \cdot B_1') = 1$

$Q(C_2 / A_2 \cdot C_3) = 0,8$ $\qquad$ $Q'(C_1' / A_1' \cdot B_1') = 1$

/A képletben szereplő, de itt nem jelzett valószinüségek értéke 0./

A felvett információt a

$$\phi_{\mathcal{A}\times\mathcal{B}\times\mathcal{C}}(\mathcal{I}) = \phi_{\mathcal{A}}(\mathcal{I}) + \phi_{\mathcal{B}}(\mathcal{I}/\mathcal{A}) + \phi_{\mathcal{C}}(\mathcal{I}/\mathcal{A}\times\mathcal{B})$$

formula alapján számoljuk, ahol a részleteredményeket a 113.oldalon levő képlet alapján kapjuk.

A számitásokat elvégezve a következő eredményekhez jutunk:

$$\phi_{\mathcal{A}}(\mathcal{I}) = 0,585$$
$$\phi_{\mathcal{B}}(\mathcal{I}/\mathcal{A}) = 0,415$$
$$\phi_{\mathcal{C}}(\mathcal{I}/\mathcal{A}\times\mathcal{B}) = 0,322$$
$$\phi_{\mathcal{A}\times\mathcal{B}\times\mathcal{C}}(\mathcal{I}) = 1,322.$$

E példa legfontosabb tanulságát akkor vonhatjuk le, ha észrevesszük, hogy ha a tanár több négyzetgyökvonást ir fel meghatározásra, azaz $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ eseményrendszerek mellé ugyanilyen tipusu eseményrendszereket veszünk, a felvett információ értéke nem változik lényegesen. Ezt abból láthatjuk, hogy a feltételes szubjektiv valószinüségek növekednek.

$$Q(A_2) = 0,4$$
$$Q(B_2 / A_2) = 0,6$$
$$Q(C_2 / A_2 \cdot C_3) = 0,8$$

Ha további példák helyes válaszát közöljük a tanulóval, várható, hogy felismeri a szabályt, tehát a helyes válaszok szubjektiv valószinüsége meg fog egyezni a matematikaival, aminek értéke 1, $\frac{1}{1}$ logaritmusa pedig 0. Ezzel olyan példát mutattunk, ahol végtelen sok /kontinuum számosságu is lehet/ eseményrendszerre vonatkozó információérték véges.

Az információérték relativitásáról még szólni fogunk. Most a négyzetgyök fogalmához 1,322 bit értékü információ alapján jutottunk. Elképzelhető /például a definició egyszerü közlése utján/, hogy kisebb értékü információval is eljuthattunk volna ilyen szintü ismerethez, más /persze fontos/ kérdés, hogy melyik információ nyujt mélyebb, tartósabb ismeretet.

A következőkben egy olyan példát irunk le, ahol az összeg nem konvergens.

Vegyünk egy s hosszuságu intervallumot, ezen véletlenszerüen jelöljünk ki egy pontot. Osszuk az adott intervallumot 2; 4; ...; 2^n; ... egyenlő részekre. Ezek a beosztások egy-egy eseményrendszert határoznak meg, ha minden esetben azt kérdezzük, hogy mely részintervallumba fog esni a pont. Ezek az eseményrendszerek egy $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n, \dots$ végtelen sorozatot alkotnak. Egyenletes matematikai és szubjektiv eloszlást feltételezve a kezdeti szubjektiv entrópiák értéke rendre 1; 2; ...; n; ...

Tegyük fel, hogy a pont megjelenése után a kisérleti személy kezébe veheti az adott intervallumot és "pontosan" megmérheti a pont helyét. /A már szerepelt hasonló példa esetében csak bizonyos távolságból becsülhette a kisérleti személy a pont helyét./

Ez a mérés a kisérleti személy számára egy közlés, melynek van informáciértéke az egyes eseményrendszerekre vonatkozóan. Mégpedig

$$\emptyset_{\mathcal{A}_1}(\mathcal{I})=1\ ,\ \emptyset_{\mathcal{A}_2}(\mathcal{I})=2\ ,\ \ldots\ ,\ \emptyset_{\mathcal{A}_n}(\mathcal{I})=n\ ,\ \ldots$$

Határozzuk meg az információ értékét az összes eseményrendszerre vonatkozóan.

Már láttuk, hogy $\emptyset_{\mathcal{A}_1}(\mathcal{I}) = 1$.
$\emptyset_{\mathcal{A}_2}(\mathcal{I}/\mathcal{A}_1)$ értéke szintén 1, mert négy intervallum közül kell ugyan eredetileg eldönteni a pont helyét, de $\mathcal{A}_1$ közlésével már csak kettő lehetőség marad, tehát I már csak ezt a log 2 = 1 bizonytalanságot szünteti meg. Ez a $\emptyset_{\mathcal{A}_2}(\mathcal{I}/\mathcal{A}_1)$-et definiáló képlet alapján is bizonyitható.

Hasonló a helyzet a következő $\emptyset_{\mathcal{A}_3}(\mathcal{I}/\mathcal{A}_1\times\mathcal{A}_2)$ meghatározásnál. Az $\mathcal{A}_3$ eseményrendszert nyolc elemi esemény alkotja, ha közöljük $\mathcal{A}_1$ és $\mathcal{A}_2$ együttes bekövetkeztét /persze most elegendő csak $\mathcal{A}_2$ kimenetelét közölni, ezzel már $\mathcal{A}_1$ kimenetelét is közöltük/ akkor megint csak két részintervallum marad a kisérleti személy számára. I közlése mondja meg, hogy a fennmaradt két intervallum közül melyikbe esett tulajdonképpen a pont. Tehát $\mathcal{A}_1\times\mathcal{A}_2$ közlése után I megint 1 bit értékü információt ad.

Ugyanilyen gondolatmenettel láthatjuk be, hogy az I értékét definiáló összeg minden tagja 1, tehát a számsor nem konvergens, I-nek nincs értéke e végtelen sok eseményrendszerre vonatkoztatva.

VIII. Az információ abszolutértéke

Eddigiekben egy közlés, vagy észlelés informácioértékét mindig adott eseményrendszerre, vagy eseményrendszerek egy halmazára értelmeztük. Több példát láttunk ezek kiszámitására, amelyekből az is kitünt, hogy egy adott I értéke lényegesen függ attól, hogy melyik eseményrendszerre vonatkoztatjuk. Ugyanannak a közlésnek információtartalma az egyik eseményrendszerre vonatkoztatva lehet 0, ugyanakkor egy másikra vonatkoztatva ettől különböző érték, esetleg egy harmadikra végtelen, azaz nincs értéke, holott a szubjektum mindhárom esetben ugyanaz.

Nem volt tehát még értelme a következő tipusu feladatnak: Adott egy I információ és egy szubjektum, határozzuk meg I értékét erre a szubjektumra. Pedig az ilyen tipusu feladat megoldhatósága szükségszerü, mitsem ér elméletünk e probléma kikerülésével.

E feladat megoldása a fentiek alapján már igen egyszerü.

Vegyünk tehát gondolatban egy I információt és egy szubjektumot. Határozzuk meg mindazon eseményrendszerek halmazát, melyeknek szubjektiv entrópiáját I megváltoztathatja az adott szubjektumban. Ha ezen eseményrendszerek halmaza véges, vagy megszámlálhatóan végtelen, akkor I abszolutértékét az adott szubjektumra vonatkozóan a

$$\Phi(I) = \Phi_{\mathcal{A}_1}(I) + \Phi_{\mathcal{A}_2}(I/\mathcal{A}_1) + \Phi_{\mathcal{A}_3}(I/\mathcal{A}_1 \times \mathcal{A}_2) + \cdots$$

összeg adja, ahol $\{\mathcal{A}_n\}$ az eseményrendszerek halmaza.

Ha $\{\mathcal{A}\}$ kontinuum számosságu, akkor megfelelő reprezentánsok kiválasztására kényszerülünk. Például ha I egyenletek megoldására ad információt, akkor nem vehetjük az összes egyenletet, de nincs is rá szükség. Ilyen esetben ki kell választanunk néhány "tipus" egyenletet, ezek megoldásából már következtethetünk a felvett információ nagyságára. Ha ugyanis egy tanuló meg tudja oldani a lg 2 x + lg /x-1/ = 5 egyenletet, akkor felesleges lenne utána nézni, hogy a lg 3 x + lg /x + 1 / = 8 egyenletet meg tudja-e oldani. Itt hasonló a helyzet, mint a négyzetgyök fogalmára hozott példánk esetében, itt is 0-vá válnak bizonyos feltételes valószinüségek.

A feltételeknek eleget tevő eseményrendszerek halmazának feltérképezése esetenként igen könnyen megtörténhet, más esetekben viszont komoly nehézségekbe ütközhetünk. Véleményünk, hogy az első esettel találkozunk sürübben. Ha egy pszichológiai kisérletnél a szubjektum észlelése által felvett információt akarjuk mérni, a kisérletvezető általában előre tudja /hisz ő tervezi meg a kisérletet/, az eseményrendszerek mely halmazára fog hatni a felvett információ. Talán még fokozottabban igaz ez a pedagógiai gyakorlatban, vagy kisérleteknél. A tanár /a diákok ismeretei alapján/ előre tudja, hogy egy adott közlés milyen hatással lesz a diákokra, tehát e közlésnek, mint információnak az értéke elvileg könnyen meghatározható. Más kérdés a szükséges mérések és számitások munkaigényessége. E tekintetben fontos, hogy a számolás könnyen programozható, igy a sok adat alapján történő számolás nem okozhat problémát a számitógépek korszakában.

Mindenesetre megmutattunk egy utat arra, hogy hogyan lehet egzaktan és a szubjektum sajátosságait figyelembevéve igen általános feltételek mellett alkalmas módon mérni az ember által felvett információ értékét.

Persze még igen sok problémát lehetne felvetni az általunk bevezetett információértékkel kapcsolatban. Ezek közül egy-két kérdést még mi is felvetünk.

Egyik ilyen probléma, hogy a kisérletező pszichológust, a kutató vagy gyakorló pedagógust nem mindig érdekli az összes eseményrendszer, melynek szubjektiv valószinüségeloszlását I megváltoztatja. Ekkor I értékét csak az érdeklődéskörbe eső eseményrendszerek halmazára számitják ki. Ha az érdeklődési kört E-vel, akkor az erre vonatkozó információértéket $\emptyset_E(\gamma)$-vel jelöljük.

Például az 1848-as szabadságharc tárgyalása után az érdekli a felmérőt, hogy e tananyag /mint információ/ milyen hatással volt a diákok internacionalista érzelmére. Ez az anyagrész ugyanis bizonyára kihat a diákok világnézetére. Ebben az esetben a felmérőnek meg kell keresnie azokat az eseményrendszereket, amelyek alapján reális képet kap a felmérés során.

Egy ilyen eseményrendszer például a következő:

A magyar-lengyel barátság alapja

A_1 : a két ország kis távolsága

A_2 : gazdasági kapcsolatok

A_3 : a történelmi multban keresendő

A_4 : egyéb okok

Valószinü, hogy az A_3 esemény szubjektiv valószinüsége nő a szabadságharc eseményeinek ismertetése után. A tanáron mulik a felvett információ nagysága erre az eseményrendszerre vonatkozóan. Ha például a tanár a kelleténél jobban kiemeli Bem hőstetteit, a külföldi segitséget, A_3 szubjektiv valószinüsége annyira megnőhet, hogy eredményül negativ információérték adódik.

Olyan eset is előfordulhat, hogy az információérték kiszámitásakor egyes eseményrendszerek kevésbé, mások jobban érdekelnek bennünket. Ebben az esetben célszerü az egyes eseményrendszerekre eső információértékeket érdeklődésünknek megfelelő sullyal figyelembe venni. Ez a felfogás az előbbinek általánositása, ugyanis ott azt tettük, hogy azokat az eseményrendszereket, melyek nem érdekelnek bennünket az információfelvétel mérésénél 0 sullyal, a számunkra fontosakat pedig 1 sullyal vettük figyelembe. Általános esetben ezektől különböző sulyokkal is számolhatunk.

Az előbbi fejezetben szerepelt az a példa, hogy egy /0; S/ intervallumon véletlenszerüen megjelenik egy pont, majd a pont pontos helyét közöljük a szubjektummal. Megállapitottuk, hogy az ott leirt eseményrendszerek halmazára nézve ennek az információnak végtelen nagy az értéke.

Más eredményt kapunk, ha ezeket az eseményrendszereket például 1; $\frac{1}{2}$; ... $\frac{1}{2^{n-1}}$; ... sulyokkal látjuk el. Ekkor a

$$\phi_{\mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_n \times \cdots}(\mathcal{I}) = 1 \cdot \phi_{\mathcal{A}_1}(\mathcal{I}) + \frac{1}{2}\phi_{\mathcal{A}_2}(\mathcal{I}/\mathcal{A}_1) + \\ + \cdots + \frac{1}{2^{n-1}}\phi_{\mathcal{A}_n}(\mathcal{I}/\mathcal{A}_1 \times \cdots \times \mathcal{A}_{n-1}) + \cdots$$

számsor már konvergens lesz, mégpedig az összeg értéke 2.

/Itt egy egyszerü $a_1 = 1$; $q = \frac{1}{2}$ mértani sorról van szó./ Tehát erre az eseményrendszer halmazra, ilyen érdeklődési feltételek mellett a közlés 2 bit információértékkel bir.

Ilyen problémával állnánk szembe, ha a tanitás során felvett információt kellene mérni. Más sullyal kellene számitanunk a törzsanyagot és megint mással a kiegészitő anyagot.

E kérdéskör vizsgálata ujabb problémákat vetne fel /ilyen probléma lenne például a kommutativitás kérdése/, részletekbe nem bocsátkozhatunk.

Megjegyezzük még, hogy $\phi(\gamma)$ értéket egyetlen szubjektumra vonatkoztattuk. Előfordulhat azonban, például a tanitási gyakorlatban, hogy I értékét szubjektumok egy halmazára kell számolnunk /pl. egy osztályra/. A számolást illetően nincs különösebb eltérés a fentiektől. Fel kell mérni az adott közösség átlagos szubjektiv eloszlását, illetve ennek megváltozását. Ennek legcélszerübb utja az egyes embereken felmért szubjektiv valószinüségek átlagának kiszámitása /ld. [16.] /. A felmért átlagos szubjektiv entrópiaváltozás alapján számitható a közösségre vonatkoztatott információérték. Az igy kapott eredmény bizonyos átlagérték, ami akkor ad reálisabb képet, ha megközelitően azonos intelligenciaszintü embereken végeztük a felmérést, ekkor ugyanis az egyes emberek által felvett információértékeknek kisebb a szórása.

IX. Információk algebrája

Az eddigiektől eltérően most több információt, ezek kapcsolatát vizsgáljuk.

Vegyük az információk egy halmazát, e halmaz bizonyos feltételek mellett algebrai strukturát alkot. E struktura leirása szintézise e dolgozatnak, reméljük, hogy ezzel az információ fogalmának további mélyitését érjük el.

A hétköznapi értelemben vett információk halmazán lehetetlen lenne eldönteni, hogy két információ egyenlő vagy sem. Mi azonban egzakt módon definiáltuk az információ fogalmát, igy e kérdés egyértelmüen eldönthető. Vigyázni kell azonban az információ relativ jellegére. Láttuk például, hogy ugyanaz a közlés nem biztos, hogy ugyanazt az információt jelenti két különböző szubjektum esetében.

Rögzitenünk kell tehát a szubjektumot és egy eseményrendszer halmazt. Ez a rögzités nagyon fontos, mert ezután /bár nem mondjuk mindig/ adott szubjektum és eseményrendszer halmaz feletti algebráról lesz szó. Ha ezek közül bármi megváltozik, ujabb algebrát kell értelmeznünk.

Információfelvételnek neveztük azt a jelenséget, amikor a szubjektumban valamely eseményrendszer szubjektiv eloszlása megváltozik. Az információt I-vel jelöltük.

Két információt egyenlőnek tekintünk $/I_1 = I_2/$, ha mindkét információ ugyanazokat az eloszlásváltozásokat hozná létre az adott szubjektumban minden felvett eseményrendszeren.

A feltételes módnak lényeges szerepe van.

Két információ egyenlőségét ugyanis nem dönthetjük el ugy, hogy egymás után közöljük a két információt, és megnézzük, hogy ugyanazt a változást hozták-e létre. Alapvető probléma, hogy az egyik információ közlése már megváltoztatja a szubjektumot, tehát a második információt már nem ugyanazzal a szubjektummal közöljük, holott ezt rögzitettnek tételeztük fel.

Annak kisérleti eldöntése tehát, hogy két információ azonos -e, két azonosnak tekinthető szubjektummal történhet. Az egyik szubjektummal I_1-et, a másikkal I_2-t közöljük és ha ugyanazt a változást állapitjuk meg, akkor a két információt egyenlőnek tekinthetjük.

Egy I információ ellentettjét $\bar{I}$-al jelöljük.

Például:

I: Az asztal zöld

$\bar{I}$: Az asztal nem zöld

/Vigyázat, I-nek nem ellentettje az az információ, hogy "Az asztal sárga"./

Azt mondjuk, hogy I_1 tartalmazza I_2-t, ha I_1 közlése után I_2 már nem hozna létre változást a szubjektumban. /A feltételes mód jelentőségét később fogjuk látni. /Ezt a kapcsolatot $I_2 \subset I_1$ - el jelöljük. Előfordulhat speciális esetként, hogy $I_2 \subset I_1$, mellett még az $I_2 = I_1$ feltétel is teljesül, ezt a kapcsolatot $I_2 \subseteq I_1$ -el jelöljük.

Például:

I_1 : Ez az ember alacsonyabb 180 cm-nél | $I_2 \subset I_1$,

I_2 : Ez az ember alacsonyabb 190 cm-nél | de $I_1 = I_2$

Üres információnak nevezzük azt az információt, mely a rögzitett szubjektumban, az adott eseményrendszer halmaz esetében nem hoz létre szubjektiv valószinüségeloszlás-változást.

Az $I_2 \subset I_1$ definicióra visszatérve a feltételes mód jelentősége most látszik, ugyanis az a kijelentés, hogy I_2 nem hoz létre szubjektiv valószinüségeloszlásváltozást, nem jelenti azt, hogy ez üres információ, hisz ez csak I_1 közlése után történne meg, ami megint a szubjektum megváltozását eredményezné. I_2 tehát már egy másik /I_1 közlése után kialakult/ szubjektumra nézve üres információ.

Az információk halmazán müveleteket is bevezethetünk. Két halmaz összegén értjük azt az $I_1 + I_2$-vel jelöllt információt, amely I_1 és I_2 egymás utáni, vagy egyidejü közlésével jön létre.

E müveletnél figyelembe kell venni a felejtés folyamatát. Világos, hogy a felejtés mint negativ információ jelentkezik a szubjektumnál. I_1 közlése után tehát az idő mulásával ennek értéke csökkenhet, illetve az általa létrehozott szubjektiv eloszlások változhatnak, ezért az összeadás müvelete csak ugy lesz egyértelmü, ha I_1 közlését I_2 olyan időn belül követi, amikor még az I_1 által létrehozott szubjektiv eloszlások nem változtak meg.

Vegyünk hat dobozt, mindegyikbe tegyünk egy golyót, a hat közül 1 piros, a többi fehér. A kisérleti személy feladata kitalálni, hogy melyik dobozba került a piros golyó.

Tekintsük a következő információkat:

I_1 : "Nincs az első kettőben"

I_2 : "Nincs az utolsó háromban"

I_1+I_2 : "A harmadikban van".

$\mathcal{A}$ jelentse a következő eseményrendszert:

{az elsőben van; ... ; a hatodikban van}, akkor

$$\phi_{\mathcal{A}}(I_1) = \log 6 - \log 4 = \log \frac{3}{2}$$

$$\phi_{\mathcal{A}}(I_2) = \log 6 - \log 3 = \log 2 = 1$$

$$\phi_{\mathcal{A}}(I_1 + I_2) = \log 6 - \log 1 = \log 6$$

Ha ezeket az eredményeket összevetjük, akkor azt tapasztaljuk, hogy általában

$$\phi_{\mathcal{A}}(I_1 + I_2) \neq \phi_{\mathcal{A}}(I_1) + \phi_{\mathcal{A}}(I_2)$$

Ebből is látható, hogy az információk halmazán értelmezett függvény nem mérték, ezért használtuk következetesen az "érték" kifejezést.

Tulajdonképpen már használtuk az I_2/I_1 jelölést. Ez azt az információt jelenti, amit I_2 ad még I_1 közlése után. Ezt az információt az I_2 és I_1 különbségének is nevezhetnénk.

Ritkábban fordul elő, de a teljesség szempontjából fontos, amikor " I_1 vagy I_2 " információt közöljük a szubjektummal, ezt $I_1 \cdot I_2$ -vel jelöljük.

Például

I_1 : Az asztal zöld

I_2 : Az asztal sárga

$I_1 . I_2$: Az asztal zöld vagy sárga.

Látható, hogy $I_1 \cdot I_2$ általában kisebb információt jelent, mint I_1, vagy I_2 külön-külön, tehát

$$I_1 \cdot I_2 \leq I_1 \qquad \text{és} \qquad I_1 \cdot I_2 \leq I_2$$

Hasonló módon könnyen belátható viszont, hogy

$$I_1 + I_2 \supseteq I_1 \quad \text{és} \quad I_1 + I_2 \supseteq I_2$$

Nyilvánvaló, hogy $I \cdot \bar{I}$ üres információt ad minden I-re. $I + \bar{I}$ viszont lehetetlen információ, ez a másik kitüntetett eleme az információk halmazának.

Ha I információ nem bontható fel valódi módon két információ szorzatára, vagyis, ha $I = I_1 \cdot I_2$ nem áll fenn, ha $I_1 \neq I$ és $I_2 \neq I$, akkor I-t elemi információnak nevezzük.

Vegyük ezek után az információk egy olyan /$\mathcal{I}$/ halmazát, hogy minden $I \in \mathcal{I}$ esetén $\bar{I} \in \mathcal{I}$ is teljesüljön. Teljesüljenek továbbá a következő feltételek:

Ha $I_1 \in \mathcal{I}$ és $I_2 \in \mathcal{I}$, akkor

$$I_1 + I_2 \in \mathcal{I} \quad \text{és} \quad I_1 \cdot I_2 \in \mathcal{I} .$$

Könnyen belátható, hogy az információk egy ilyen halmaza Boole-algebrát, illetve, ha megszámlálhatóan végtelen sok eleme van $\mathcal{I}$ -nek, akkor Borel-algebrát alkot. Pontosabban az adott szubjektum és a rögzitett eseményrendszer halmaz feletti algebráról van szó.

Ezen algebrán $\emptyset$ egy függvényt ad meg, ugyanis minden információhoz egy egyértelmüen meghatározott értéket rendel.

I R O D A L O M

1. dr.ADORJÁN CS.: A valószinüségi modellezés néhány kérdéséről az invariáns mozzanatok percepciójában.
Pszichológiai tanulmányok X. 1967.

2. ÁDÁM GY.: Érzékelés tudat emlékezés
Gondolat Budapest 1976.

3. ARATÓ M.: A matematika hazai alkalmazásának helyzete
Magyar Tudomány 1977/7-8.

4. BÉKÉS F.: Ismeretmérés, ismeretstruktura, ismerettipológia
Tömegkommunikációs Kutatóközpont Budapest 1980.

5. FENYŐ I. - FREY T.: Matematika villamosmérnököknek I.
Müszaki Könyvkiadó Budapest 1964.

6. C. GÜNTER.: Az információelmélet alkalmazása tanuláslélektani problémák megoldásánál
Tanulmányok a programozott tanitás köréből 1973.

7. HÓDOS T.: Müszaki pszichológia
Alkalmazott pszichológia
Gondolat 1968.

8. A.M.JAGLOM, I.M. JAGLOM, A.JA.MINCSIN
Az információelmélet matematikai alapjai
Budapest 1959.

9. KLEIN S.: Az információelsajátitás és alkotó alkalmazás /JA/ mérése oktatógép segitségével.
Magyar Pszichológiai Szemle 1972.

10. B.F. LOMOV: Ember és technika. A müszaki pszichológia alapjai
Akadémiai Kiadó. Budapest 1969.

11. B.F. LOMOV: A müszaki pszichológia perspektivikus problémái
Pszichológiai tudományok X. 1967.

12. I.P. NATANSZON: Konstruktiv függvénytan
Akadémiai Kiadó 1952.

13. PRÉKOPA A.: Valószinüségelmélet
Müszaki Könyvkiadó 1974.

14. Dr.RÓKUSFALVY P.: A teljesitmény motiváció és a döntés információelméleti vonatkozásai
Ergonómia 1971/3.

15. SZEKERES I.: Egy speciális tanulási folyamat vizsgálata információelméleti módszerekkel.
Magyar Pszichológiai Szemle 1976/5.

16. SZEKERES I.: Szubjektiv valószinüségeloszlás különböző tipusu feladathelyzetekben
Magyar Pszichológiai Szemle 1979/3.

17. OLEG A. KONOPKIN: Az információbefogadás és az emberi tevékenység szándékos, tudatos jellege.
A pszichikum és a tevékenység a mai szovjet pszichológiában 1974.

18. TONI J.: Néhány informativ értékelés egy pszichológiai kisérletről

19. VASZKÓ M.: Ipari nagyrendszerek operátori tevékenységének információelméleti vizsgálata
Ergonómia 1971/1.

20. VASZKÓ M.: Munkalélektan
Tankönyvkiadó, Budapest 1970.

21. VINCE I.: Az információelmélet egy fogalmának értelmezéséről
Matematikai Lapok 1959.

22. ITELSZON: Matematikai és kibernetikai módszerek a pedagógiában
Tankönyvkiadó, Budapest 1967.

23. L.N. LANDA: Algoritmizálás az oktatásban
Tankönyvkiadó, Budapest 1969.

KÜLÖNLENYOMAT

MATEMATIKAI LAPOK

23. ÉVFOLYAM 3—4. SZÁMÁBÓL

Szekeres István

A tanulás folyamatának függvénytani leírása

BOLYAI JÁNOS MATEMATIKAI TÁRSULAT BUDAPEST, 1972

A TANULÁS FOLYAMATÁNAK FÜGGVÉNYTANI LEÍRÁSA

SZEKERES ISTVÁN

E dolgozat az emberi tanulást az idő függvényében vizsgálja matematikai módszerrel, ezzel a tanulás folyamatának egy matematikai modelljét adja meg.

A tanulás igen általános és bonyolult pszichológiai fogalom, ezért nem törekedhetünk annak teljes matematikai leírására, csupán a felejtést és a reminiszcenciát vesszük figyelembe. Modellünket azonban úgy építjük fel, hogy további tanulást meghatározó tényezők (fáradás, a tanuló és a tanulandó anyag viszonyának változása stb.) figyelembevételére is legyen lehetőség.

A pszichológiából ismert fogalmakat nem definiáljuk, azokat alapfogalomként kezeljük. Ilyen alapfogalom maga a tanulás, a tanuló, az anyag (a tanulás tárgya), a tanuló pszichológiai állapota, stb.

Megállapításainkat egyetlen tetszőlegesen rögzített tanulóra vonatkoztatjuk, akinek pszichológiai állapota a tanulás megkezdésekor adott, majd a tanulás folyamán az általunk meghatározott módon változik. Ugyancsak tetszőlegesen rögzített a megtanulandó anyag.

Fontos feladatunk, hogy az anyagot mérhetővé tegyük.

Egységnyinek vesszük azt az anyagmennyiséget, melyet a tanuló egységnyi idő alatt tud megtanulni a nulla időpillanathoz tartozó pszichológiai állapot mellett. (A nulla időpillanat a tanulás megkezdésének pillanata, ahonnan az időt fogjuk számítani.)

Az anyag egysége tehát függ a tanulótól és annak kezdeti pszichológiai állapotától, amiket mi rögzítettünk. Definíciónk alapján a tanuló számára „nehezebb" anyagot mi nem nehezebbnek, hanem többnek vesszük.

E bevezető után áttérünk a modell leírására, ahol mindenek előtt a tanulás sebességét definiáljuk.

1. A TANULÁS SEBESSÉGE

Egy adott pillanatban a tanuló tanulási sebességét azzal a maximális anyagmennyiséggel mérjük, melyet a tanuló az adott pillanathoz tartozó pszichológiai állapot mellett egységnyi idő alatt tud áttanulmányozni úgy, hogy ha közben nem felejt, akkor az így áttanulmányozott anyagot hiánytalanul produkálni tudja.

Az anyag egységének definíciójából következik, hogy a tanulás sebessége a tanulás megkezdésekor egységnyi. A tanulás folyamán a tanulás sebessége állandóan változik, aminek oka sokrétű. A tanulás sebességét csökkentő tényező például a fáradás. A tanulás folyamán megváltozik a tanuló és az anyag viszonya (mi egyszerűen pszichológiai állapotváltozásnak tekintjük), ami a sebességet növelő tényezőként szerepelhet.

A tanulás sebességét x időpillanatban $\lambda(x)$-szel jelöljük és feltesszük, hogy $\lambda(x)$ rendelkezik a következő tulajdonságokkal:

a) $\lambda(x)$ értelmezési tartománya a folyamatos tanulás intervalluma,

b) $\lambda(0)=1$,

c) $\lambda(x)>0$,

d) $\lambda(x)$ az értelmezési tartomány minden pontjában folytonos.

Amíg más feltételt nem teszünk mindig folyamatos tanulással dolgozunk, a tanulás megszakítása után ugyanis a sebességfüggvény megváltozik, amit könnyen beláthatunk, ha a pihenés folyamatára gondolunk.

Vegyünk fel a folyamatos tanulás intervallumán egy $[x, x+\Delta x]$ intervallumot és tekintsük ezen a tanulás sebességét egy állandó $\lambda(\xi)$ értéknek, ahol $x \leqq \xi \leqq x+\Delta x$. Ha $T(x)$-szel jelöljük a $[0, x]$ intervallumon áttanulmányozott anyagmennyiséget, akkor az előbbiek szerint igaz a

$$T(x+\Delta x)-T(x) = \lambda(\xi)\cdot\Delta x, \quad \text{ahol} \quad x<\xi\leqq x+\Delta x$$

egyenlőség, ahonnan

$$T'(x) = \lambda(x), \tag{1.1}$$

illetve a $T(0)=0$ feltételt felhasználva

$$T(x) = \int_0^x \lambda(t)\,dt. \tag{1.2}$$

Szükségünk lesz még az $[x_1, x_2]$ intervallumon áttanulmányozott anyagmennyiségre, melyet

$$T(x_2)-T(x_1) = \int_{x_1}^{x_2} \lambda(t)\,\mathrm{d}t \tag{1.3}$$

formában írhatunk fel.

2. A TANULÁS ÁLTALÁNOS EGYENLETEI

Vegyünk fel egy tetszőleges nagyságú a anyagmennyiséget, melyről tegyük fel, hogy a tanuló a tanulás megkezdése után t_1 idő elteltével egyetlen pillanatban tanulta. Egy t_2 ($\geqq t_1$) időpillanatban már nem a anyagmennyiséget tud az előbb tanultból, hanem valamely $a\cdot D(t_1, t_2-t_1)$-t, azaz egy t_1-től és t_2-től függő számmal szorzódik. E változás oka speciálisan lehet a felejtés, melyet a következő fejezetben részletesen vizsgálunk.

A most bevezetett $D(x_1, x_2)$ függvényre az alábbi feltételeket tesszük:

a) $D(x_1, 0)=1$,

b) $1\geqq D(x_1, x_2)>0$,

c) az első változó folytonos, a második változó differenciálható függvénye.

A továbbiakban a folyamatos tanulás vizsgálatánál mindig a most bevezetett $D(x_1, x_2)$ függvény leírása lesz a feladatunk, mert ezzel felírhatjuk az x-ben tudott $f(x)$ anyagmennyiséget, ha a tanuló a $[0, x]$ intervallumon folyamatosan tanul az első fejezetben leírt sebességgel, de bármely pillanatban tanult anyag $D(x_1, x_2)$ függvény szerint változik.

$f(x)$ meghatározása érdekében vegyük a $[0, x]$ intervallum egy beosztását, ahol az i-edik osztáspontot x_i-vel jelöljük. Vizsgáljuk először az $[x_i, x_{i+1}]$ intervallumon tanult anyag változását.

Az $[x_i, x_{i+1}]$ intervallumon tanult anyagmennyiség (1.3) alapján

$$T(x_{i+1})-T(x_i) = \int_{x_i}^{x_{i+1}} \lambda(t)\,dt,$$

vagy

$$T(x_{i+1})-T(x_i) = \lambda(\xi_i)\cdot(x_{i+1}-x_i), \quad \text{ahol} \quad x_i \leqq \xi_i \leqq x_{i+1}.$$

Tekintsük ezt az anyagmennyiséget egy adott ξ_i pillanatban tanult anyagmennyiségnek. Mivel ennek az x-ben tudott részét keressük, ezért $x-\xi_i$ ideig változik, tehát $D(\xi_i, X-\xi_i)$-vel szorzódik. Az $[x_i, x_{i+1}]$ intervallumon tanult anyag x-ben tudott része tehát

$$\lambda(\xi_i)\cdot(x_{i+1}-x_i)\cdot D(\xi_i, x-\xi_i), \quad \text{ahol} \quad x_i \leqq \xi_i \leqq x_{i+1}.$$

Az egyes részintervallumokban tanult anyagok x-ben tudott részeit összegezve kapjuk az x-ben tudott összes anyagot:

$$f(x) = \sum_{i=1}^{n} \lambda(\xi_i)\cdot D(\xi_i, x-\xi_i)\cdot(x_{i+1}-x_i), \quad \text{ahol} \quad x_i \leqq \xi_i \leqq x_{i+1}.$$

Mivel a jobb oldal folytonos függvény integrálközelítő összege, ezért

$$f(x) = \int_0^x \lambda(t)\cdot D(t, x-t)\,dt. \tag{2.1}$$

Ehhez hasonlóan adjuk meg az x-ben tudott anyagmennyiséget abban az esetben is, ha a tanuló egy $y(\leqq x)$ időpillanatban abbahagyta a tanulást.

Tegyük fel először, hogy a tanulás ideje alatt és a tanulás abbahagyása után ugyanazon $D(x_1, x_2)$ függvény szerint változik egy adott pillanatban tanult anyag tudott része. Ezt figyelembe véve megismételhetnénk az előbbi gondolatmenetet, csupán az integrációs tartomány változik, hisz a $[0, y]$ intervallumon tanult folyamatosan a tanuló. Ezért, ha $F(x, y)$-nal jelöljük az y-ig tanultból x-ben tudott anyagmennyiséget, akkor

$$F(x, y) = \int_0^y \lambda(t)\cdot D(t, x-t)\,dt \tag{2.2}$$

összefüggéshez jutunk.

Még általánosabb összefüggést kapunk, ha nem tesszük azt a nagyon is önkényes megszorítást, hogy a tanulás abbahagyása után ugyanúgy változik a tudott anyag, mint a tanulás ideje alatt.

Tegyük fel tehát, hogy a tanulás ideje alatt továbbra is $D(x_1, x_2)$ függvény szerint változik egy adott pillanatban tanult anyag tudott része, a tanulás abbahagyása után pedig egy $B(x_1, x_2)$ függvény írja le a változást, ahol az első változó az anyag megtanulásától a tanulás abbahagyásáig eltelt időt, a második változó pedig a tanulás abbahagyásától eltelt időt jelenti. Pontosabban egy t pillanatban tanult a anyag változása a következő:

a) y-ban (a tanulás abbahagyásakor) tudott anyagmennyiség:

$$a \cdot D(t, y-t),$$

b) $x(>y)$-ben tudott anyagmennyiség:

$$a \cdot D(t, y-t) \cdot B(y-t, x-y).$$

A most bevezetett $B(x_1, x_2)$ függvényről feltesszük, hogy $B(x_1, 0)=1$ minden x_1-re, és $B(x_1, x_2)$ mindkét változó folytonos függvénye.

Készítsük el a $[0, y]$ intervallum egy beosztását, majd a fentieket figyelembe véve állítsuk elő $F(x, y)$-t integrálközelítő összeg formájában. A már bemutatott gondolatmenet alapján kapjuk, hogy

$$F(x, y) = \sum_{i=1}^{n} \lambda(\xi_i) \cdot D(\xi_i, y-\xi_i) \cdot B(y-\xi_i, x-y) \cdot (y_{i+1}-y_i),$$

ahol

$$y_i \leqq \xi_i \leqq y_{i+1}.$$

A beosztást minden határon túl finomítva

$$F(x, y) = \int_0^y \lambda(t) \cdot D(t, y-t) \cdot B(y-t, x-y)\, dt \tag{2.3}$$

összefüggéshez jutunk.

Mivel eddig a tanulás folyamatára semmiféle megszorítást nem tettünk, ezért a fenti összefüggéseket joggal nevezhetjük a tanulás általános egyenleteinek.

A következő fejezetekben részletesebben vizsgáljuk a felejtést és a reminiszcenciát, mivel azonban kísérleti eredményekre nem támaszkodunk, ezért továbbra is általános összefüggésekhez jutunk, melyekkel módszert adunk a kísérleti eredmények vizsgálatához.

3. A FELEJTÉS VIZSGÁLATA

A továbbiakban az áttanulmányozott anyag két részre oszlik: a tudott anyagra és az elfelejtett anyagra.

Vegyünk fel most is egy tetszőleges nagyságú a anyagmennyiséget és tegyük fel, hogy a tanuló egy adott pillanatban tanulta. A felvett a anyag az idő múlásával állandóan csökken, ezt a változást az előző fejezetben bevezetett $D(x_1, x_2)$-nek megfelelő függvény írja le. A tárgyalás egyszerűsége érdekében azonban megelégszünk a következő speciális esettel: feltesszük, hogy egy adott pillanatban tanult anyag felejtése nem függ a megtanulásának időpontjától, csupán az azóta eltelt időtől, feltesszük továbbá, hogy az anyag felejtése ugyanúgy történik a tanulás ideje alatt, mint a tanulás abbahagyása után. Ezeket a feltételeket figyelembe véve látható, hogy a tudott anyag változását egy egyváltozós egyetemes függvény írja le. Jelöljük ezt a függvényt $D(x, t)=E(t)$-vel, ahol t a megtanulás óta eltelt időt jelenti, azaz a adott pillanatban tanult anyagból t idő múlva a tanuló $a \cdot E(t)$ mennyiségre emlékszik, vagyis ennyi a tudott anyagmennyiség. E jelölést alkalmazva x idejű folyamatos tanulás után az x-ben tudott $f(x)$ anyagmennyiséget (2.1) felhasználásával a következőképpen írhatjuk fel:

$$f(x) = \int_0^x \lambda(t) \cdot E(x-t)\, dt. \tag{3.1}$$

Hasonlóan módosul az $F(x, y)$ függvény is:

(3.2) $$F(x, y) = \int_0^y \lambda(t) \cdot E(x-t)\,dt.$$

További feladatunk $E(t)$ függvény leírása.

Ha azt tennénk fel, hogy a tanuló egyenlő időközönként egyenlő részét felejti az anyagnak, akkor

$$E(t) = \left(1 - \frac{1}{p}\right)^t \text{ lenne, ahol } \frac{1}{p} \text{ a felejtés aránya,}$$

vagyis időegység alatt a tudott anyag ennyiedrészét felejtené el a tanuló. Ez azonban nem fedné a valóságot, hisz könnyen belátható, hogy ilyen feltétel mellett állandó tanulási sebesség esetén a tanulásnak felső határa lenne, ami a gyakorlatnak ellentmond. Ezért feltesszük, hogy a tanuló a régebbi anyagot kisebb arányban felejti, mint az újabban tanultat. Ezt úgy fogjuk fel, hogy a tudott anyag az idő múlásával egyre nagyobb súllyal rendelkezik a felejtéssel szemben.

Vizsgáljuk a felvett a anyagmennyiség változását először olyan feltétel mellett, hogy annak tulajdonsága a felejtéssel szemben változatlan, vagyis olyan tulajdonságú marad, mint a tanulás pillanatában volt. Ekkor elfogadjuk azt a feltevést, hogy a tanuló egyenlő időközönként egyenlő részét felejti a tudott anyagnak és az egységnyi idő alatt elfelejtett rész arányát $\frac{1}{p}$-vel jelöljük. Röviden a nulla időpillanathoz az $\frac{1}{p}$ felejtési arányt rendeltük. Ha viszont a megtanulás után t idő elteltével követeljük meg, hogy az anyag tulajdonsága a felejtéssel szemben állandóvá váljék, akkor most is egyenlő időközönként egyenlő részét felejtené tovább, viszont egységnyi idő alatt nem $\frac{1}{p}$, hanem ennél kevesebb $\frac{1}{p\varphi(t)}$-ed részét felejtené a t-ben tudott anyagnak

$$\left(1 - \frac{1}{p\varphi(t)}\right)^r$$

-ed részére emlékezne a tanuló. A felejtés folyamatára ezeket a feltevéseket fogadjuk el a továbbiakban.

A most bevezetett $\varphi(t)$ függvényt súlyfüggvénynek nevezzük és azt mondjuk, hogy egy adott pillanatban tanult anyag tudott részének felejtéssel szembeni súlya (röviden súlya) t idő múlva $\varphi(t)$.

Feltesszük, hogy $\varphi(t)$ függvény rendelkezik a következő tulajdonságokkal:

a) Értelmezve van minden $0 \leqq t < \infty$ pontban.

b) $\varphi(0) = 1$.

c) Szigorúan monoton növő.

d) Folytonos az értelmezési tartomány minden pontjában.

Feltesszük még, hogy $p > 1$.

Ezek után áttérünk az $E(t)$ függvény meghatározására.

A felvett a anyagmennyiségből tehát a tanuló t idő múlva $a \cdot E(t)$ mennyiséget tud. Elég kicsi Δt időnövekedést feltételezve az anyag súlyát változatlannak tekinthetjük, így $t + \Delta t$-ben ennek

$$\left(1 - \frac{1}{p\varphi(t)}\right)^{\Delta t}$$

ed részét tudja a tanuló, tehát

$$a \cdot E(t+\Delta t) = a \cdot E(t) \cdot \left(1 - \frac{1}{p\varphi(t)}\right)^{\Delta t},$$

melyből

$$\frac{a \cdot E(t+\Delta t)}{a \cdot E(t)} = \left(1 - \frac{1}{p\varphi(t)}\right)^{\Delta t}.$$

Mindkét oldal természetes alapú logaritmusát véve és Δt-vel osztva az

$$\frac{\ln E(t+\Delta t) - \ln E(t)}{\Delta t} = \ln\left(1 - \frac{1}{p\varphi(t)}\right)$$

differencialhányadoshoz jutunk, melyből $\Delta t \to 0$ határátmenettel az

$$[\ln E(t)]' = \ln\left(1 - \frac{1}{p\varphi(t)}\right)$$

egyenletet kapjuk. Mivel $E(0) = 1$ (l. $D(x_1, x_2)$-re adott feltétel), ezért

$$E(t) = e^{\int_0^t \ln\left(1 - \frac{1}{p\varphi(u)}\right) du}. \tag{3.3}$$

A felejtés modelljének leírásakor alapgondolatunk volt, hogy a tudott anyaghoz egy számot rendeltünk, melyet az anyag súlyának neveztünk. Az x-ben tudott $f(x)$ anyagmennyiség súlyeloszlása nem homogén, hisz az egyes részeit különböző időben tanulta. A legrégebben (0 időpillanatban) tanult anyag súlya $\varphi(x)$, még a legfrissebb (x-ben tanult) anyag súlya 1, a többi anyag súlya e két érték között változik.

A súlyeloszlást szemléletessé tehetjük, ha felírjuk a (3.1) alatti integrál ekvidisztans osztáspontokhoz tartozó

$$f(x) = \sum_{i=1}^{n} \lambda(\xi_i) \cdot E(x - \xi_i) \cdot \Delta r$$

közelítő összegét, ahol

$$\Delta r = \frac{x}{n} \quad \text{és} \quad (i-1)\Delta r \leqq \xi_i \leqq i\Delta r.$$

A felírt összeg egy tetszőleges tagja a ξ_i valamely Δr nagyságú környezetében tanult $\lambda(\xi_i)\Delta r$ anyag tudott része x-ben, tehát súlya $\varphi(x - \xi_i)$.

A súlyeloszlás további vizsgálata érdekében vezessük be a

$$h = h(r) = f(x) - F(x, r)$$

jelölést. Mivel h az $[r, x]$ intervallumon tanult anyag x-ben tudott részét jelenti, ezért tekinthetjük a $[0, f(x)]$ intervallum változójának. Az r pontban tanult anyag súlyát nevezzük $f(x)$ h pontbeli súlyának és $\varphi_x(h)$-val jelöljük. Nyilvánvaló, hogy

$$\varphi_x(h) = \varphi(x - r).$$

Mindkét oldalt r szerint deriválva kapjuk, hogy

$$\frac{d\varphi_x(h)}{dr} = \frac{d\varphi_x(h)}{dh} \cdot \frac{dh}{dr} = -\varphi'(x-r).$$

Mivel

$$h'(r) = -\lambda(r) \cdot E(x-r),$$

ezért

$$\frac{d\varphi_x(h)}{dh} = \frac{\varphi'(x-r)}{\lambda(r) \cdot E(x-r)}. \tag{3.4}$$

4. AZ ISMÉTLÉS LEÍRÁSA

Ebben a fejezetben az előbbi gondolatmenetet visszük tovább az elfelejtett anyag újratanulásának, ismétlésének leírásával.

Tegyük fel, hogy a $[0, b]$ intervallumon folyt az előbbiekben leírt tanulás. Ez idő alatt $T(b)$ anyagmennyiséget tanulmányozott át a tanuló és ebből b-ben tud $f(b)$-t, tehát az összes elfelejtett anyag

$$T(b) - f(b) = \int_0^b \lambda(t) \cdot [1 - E(b-t)] \, dt$$

nagyságú. b-től ezt az elfelejtett anyagot kezdi újratanulni, illetve ismételni a tanuló. Már most megállapíthatjuk, hogy az átismétlendő anyag $T(b)-f(b)$-nél nagyobb lesz, mivel az ismétlés folyamán felejt a tanuló.

Mielőtt azonban a részletes vizsgálatokat elkezdenénk, írjuk le a modell pontos tulajdonságait. Feladatunk most is a sebességfüggvény és az adott pillanatban tanult anyag további változásának megadása.

Az ismétlés elkezdésétől számított t időpillanatban a tanulás sebességét jelöljük $\bar{\lambda}(t)$-vel. Mivel ezzel részletesen nem foglalkozunk, csak annyit jegyzünk meg róla, hogy $\bar{\lambda}(0) > \lambda(b)$ feltétel látszik helyesnek, mivel most már az egyszer tanult anyag újratanulásáról van szó. Ugyanakkor valószínűnek látszik, hogy a tanulás sebességét ezután a fáradás határozza meg, mert a tanuló és az anyag viszonya nem változik lényegesen.

Az átismételt anyag tudott része változásának leírásához először a súly fogalmát kell általánosítanunk. Az egy adott pillanatban elfelejtett anyaghoz is súlyt rendelünk, mely egy $\Phi(s, t)$ függvény szerint változik, ahol s az anyag súlya az elfelejtés pillanatában, t pedig az azóta eltelt idő.

$\Phi(s, t)$-re feltesszük, hogy

a) Mindkét változó folytonos függvénye.
a) t-nek szigorúan monoton csökkenő függvénye.
c) $\Phi(s, 0) = s$.
d) $\Phi(1, t) \equiv 1$.
e) $\lim\limits_{t \to \infty} \Phi(s, t) = 1$.
f) $s \geqq \Phi(s, t) \geqq 1$.

Tegyük fel, hogy a tanuló egy adott pillanatban $\Phi(s, t)$ súlyú elfelejtett anyagot ismétel. Az újratanulás pillanatától ismét a $\varphi(x)$ függvény szerint változzon az újratanult anyag tudott részének súlya. Pontosabban az újratanulástól számított x idő múlva az átismételt anyag tudott részének súlya

$$\varphi[\varphi^{-1}(\Phi(s,t))+x] \tag{4.1}$$

lesz.

(Itt kihasználtuk, hogy $\varphi(x)$ szigorúan monoton.)

Definiálni fogjuk az adott pillanatban tanult anyag t idő alatt elfelejtett részének súlyát.

Vegyünk fel egy adott pillanatban (először) tanult a anyagmennyiséget. Mivel ezt folyamatosan felejti, ezért egy bizonyos t idő múlva az elfelejtett rész súlyösszetétele változó, mi ezeknek a súlyoknak az átlagát fogjuk venni. Osszuk fel a $[0, t]$ intervallumot ekvidisztans osztáspontokkal n egyenlő részre, majd $\frac{t}{n}$-t jelöljük Δy-nal. Mivel az $[i\Delta y, (i+1)\Delta y]$ intervallumon

$$a\cdot E(i\Delta y)-a\cdot E((i+1)\Delta y)=-a\cdot[E((i+1)\Delta y)-E(i\Delta y)]$$

anyagmennyiséget felejt el a tanuló, melynek súlya az elfelejtés pillanatában $\varphi(\xi_i)$, ahol $i\Delta y \leqq \xi_i \leqq (i+1)\Delta y$, ezért az a-ból elfelejtett összes anyag t-ben

$$-a\cdot\sum_{i=0}^{n-1}[E((i+1)\Delta y)-E(i\Delta y)],$$

ahol az egyes tagok súlya rendre

$$\Phi[\varphi(\xi_i), t-\xi_i].$$

Vegyük a

$$\frac{-a\cdot\sum\limits_{i=0}^{n-1}\{[E((i+1)\Delta y)-E(i\Delta y)]\cdot\Phi[\varphi(\xi_i), t-\xi_i]\}}{-a\cdot\sum\limits_{i=0}^{n-1}[E((i+1)\Delta y)-E(i\Delta y)]}$$

súlyozott átlagot, majd ennek $n\to\infty$-ben vett határértékét. E határértéket $\psi(t)$-vel jelölve

$$\psi(t)=\frac{1}{E(t)-1}\int_0^t E'(y)\cdot\Phi[\varphi(y), t-y]\,dy \tag{4.2}$$

összefüggéshez jutunk. Az így definiált $\psi(t)$-t nevezzük egy adott pillanatban tanult anyag t idő alatt elfelejtett része súlyának.

Kezdje el a tanuló az ismétlést b-ben, mégpedig az eredeti sorrendet megtartva. Jelöljük $k(z)$-vel az eredetileg $[0, z]$ intervallumon tanult anyag elfelejtett részének ismétléséhez szükséges időt. Az egész anyag átismétlése tehát $k(b)$ ideig tart. Vegyünk fel a $[0, b]$ intervallumon egy z és egy $z+\Delta z$ pontott, ezeknek a $[0, k(b)]$ intervallumon megfelel egy $k(z)$, illetve $k(z+\Delta z)$ pont. Mivel a $[b+k(z)\ b+k(z+\Delta z)]$ intervallumon a tanuló $k(\xi)$ ideje ismétel, ahol $z\leqq\xi\leqq z+\Delta z$, ezért itt

$$\bar{\lambda}(k(\xi))\cdot[k(z+\Delta z)-k(z)]$$

anyagmennyiséget tanulmányoz át. Másrészt viszont ezen az intervallumon az eredetileg z-től $z+\Delta z$-ig tanult $\lambda(\xi)\cdot\Delta z$ anyagmennyiségből elfelejtett

$$\lambda(\xi)\cdot\Delta z-\lambda(\xi)\cdot\Delta z\cdot E[b+k(\xi)-\xi]$$

anyagmennyiséget tanulmányozza át (ugyanis $b+k(\xi)-\xi$ ideig felejtette), ezért

$$\bar{\lambda}(k(\xi))\cdot[k(z+\Delta z)-k(z)] = \lambda(\xi)\cdot\Delta z-\lambda(\xi)\cdot\Delta z\cdot E[b+k(\xi)-\xi],$$

melyből

$$\bar{\lambda}(k(z))\cdot k'(z) = \lambda(z)-\lambda(z)\cdot E[b+k(z)-z] \tag{4.3}$$

differenciálegyenletet kapjuk $k(z)$-re.

Vegyünk fel a $[b, b+k(b)]$ intervallumon egy tetszőleges x pontot és jelöljük $m_b(x)$-szel a b-től x-ig átismételt anyagnak x-ben tudott részét. $m_b(x)$ meghatározása végett vegyünk fel továbbá egy tetszőleges y és $y+\Delta y$ pontot a $[b, x]$ intervallumon, melyeknek az előbbiek szerint a $[0, b]$ intervallumon egy $k^{-1}(y-b)$ és egy $k^{-1}(y+\Delta y+b)$ pont felel meg. A felvett $[y, y+\Delta y]$ intervallumon a tanuló $\bar{\lambda}(\xi-b)\cdot\Delta y$ anyagmennyiséget tanulmányoz át, ahol $y\leqq\xi\leqq y+\Delta y$, melynek súlya az ismétléskor

$$\psi[\xi-k^{-1}(\xi-b)].$$

Mivel s súllyal induló a mennyiségű anyagból t idő múlva a tanuló

$$a\cdot\frac{E[\varphi^{-1}(s)+t]}{E[\varphi^{-1}(s)]}$$

mennyiségre emlékszik, ezért x-ben az $[y, y+\Delta y]$ intervallumon tanult anyagból

$$\bar{\lambda}(\xi-b)\cdot\Delta y\cdot\frac{E\{\varphi^{-1}[\psi(\xi-k^{-1}(\xi-b))]+x-\xi\}}{E\{\varphi^{-1}[\psi(-k^{-1}(\xi-b))]\}}$$

mennyiségre emlékszik a tanuló. Ezt minden részintervallumra meghatározva és összegezve kapjuk $m_b(x)$ egy integrálközelítő összegét, melyből határátmenettel az

$$m_b(x) = \int_b^x \bar{\lambda}(y-b)\cdot\frac{E[\varphi^{-1}\{\psi(y-k^{-1}(y-b))\}+x-y]}{E[\varphi^{-1}\{\psi(y-k^{-1}(y-b))\}]} \tag{4.4}$$

összefüggéshez jutunk.

Az összes x-ben tudott $I(b, x)$ anyagot úgy kapjuk, hogy $m_b(x)$-hez hozzáadjuk az ismétlés nélkül is tudott anyagmennyiséget, tehát

$$I(b, x) = m_b(x)+F(x, b). \tag{4.5}$$

Megjegyezzük még, hogy ha a tanuló nem kezdi meg az ismétlést a tanulás abbahagyásakor (b-ben), hanem egy ($b<$) c időpillanatban, akkor a tanulási folyamat a tanulás sebességében és a tanulandó anyag súlyában fog eltérni a fentitől, így más lesz a $k(z)$ függvény is.

Befejezésül vizsgáljuk meg az x-ben tudott anyagmennyiség szempontjából az alábbi két eset közötti különbséget:

a) A tanuló folyamatosan tanul x-ig ismétlés nélkül.

b) A tanuló b-ig tanul folyamatosan, majd x-ig ismétel, ahol $b<x\leqq b+k(b)$ feltétel teljesül.

a) esetben az x-ben tudott anyagmennyiség $f(x)$, melyet a következő formában is felírhatunk:

$$f(x) = \int_0^b \lambda(t)\cdot E(x-t)\,dt+\int_b^x \lambda(t)\cdot E(x-t)\,dt.$$

Ha viszont a b) esetet tekintjük, akkor az x-ben tudott anyagot (4.4) és (4.5) alapján

$$I(b, x) = \int_0^b \lambda(t) \cdot E(x-t)\,dt + \int_b^x \bar{\lambda}(t-b) \cdot \frac{E[\tau(t)+x-t]}{E[\tau(t)]}\,dt$$

alakban írhatjuk, ahol $\tau(t)$ jelentése (4.4)-ből leolvasható.

Vegyük e két érték különbségét:

$$I(b, x) - f(x) = \int_0^x \left\{ \bar{\lambda}(t-b) \cdot \frac{E[\tau(t)+x-t]}{E[\tau(t)]} - \lambda(t) \cdot E(x-t) \right\} dt.$$

Könnyen belátható, hogy általában igaz az

$$\frac{E(c+t)}{E(c)} > E(t) \text{ egyenlőtlenség.}$$

Ha még $\bar{\lambda}(t-b) > \lambda(t)$ (nem szükséges feltétel) is teljesül, akkor nyilvánvaló, hogy

$$I(b, x) > f(x).$$

Bizonyos feltételek teljesülésekor tehát az ismétléssel nagyobb tudásra lehet szert tenni, mint ismétlés nélkül ugyanannyi idő alatt.

5. REMINISZCENCIA VIZSGÁLATA

A pszichológiában reminiszcenciának nevezik azt a jelenséget, hogy a tanulás utáni közvetlen felidézés általában kisebb tudást mutat, mint egy bizonyos idő elteltével történő felidézés, vagyis a tanulás befejezése után egy ideig növekszik a produkálható anyag mennyisége.

x-ben produkálhatónak nevezzük tehát azt a $g(x)$ anyagmennyiséget, melyet a tanuló x-beli felidézésekor el tud mondani az x-ig tanult anyagból. A tudott (el nem felejtett) és a produkálható anyag különbségét gátlás alatt levő anyagmennyiségnek fogjuk nevezni. A reminiszcenciát úgy fogjuk értelmezni, hogy a tanulás abbahagyása után az ott gátlás alatt levő anyag „felszabadul", produkálhatóvá válik. A gátlást és a felszabadulást egy-egy folyamatként kezeljük, melyet úgy képzelünk el, hogy egy bizonyos idő alatt történő tanulás hatására gátlás alá kerül az addig produkálható anyag egy meghatározott része, majd azonnal elkezdődik (a tanulás ideje alatt is) ennek a gátlás alatt levő anyagnak a felszabadulása. Így a tudott anyag mindig két részre oszlik: a produkálható és a gátlás alatt levő anyagra. Szerepelni fog még a felszabadult anyag fogalma, mely olyan anyagot jelent, amelyik gátlás alatt levő anyagból lett produkálhatóvá, de a modellben nem különböztetjük meg, vagyis ugyanolyan produkálható anyagnak tekintjük, mint ami még nem volt gátlás alatt.

Kezdjük most is a modell pontos leírásával.

A tudott anyag felejtése továbbra is a harmadik fejezetben leírt módon történik, függetlenül attól, hogy az anyag produkálható vagy gátlás alatt levő állapotban van.

Egy adott pillanatban tanult anyag gátlás alá kerülése és felszabadulása (a felejtéshez hasonlóan) független a megtanulás idejétől.

A gátlásról feltesszük, hogy csak a tanulás ideje alatt megy végbe és mindig csak a produkálható anyagra vonatkozik. Hogy feltevésünket pontosan megtehessük, vegyünk fel egy adott pillanatban produkálható anyagmennyiséget, melynek nagysága tetszőleges a, majd tételezzük fel, hogy a tanuló tanul, de csak a gátlás érvényesül, tehát tekintsünk el a felejtéstől és a felszabadulástól. Ilyen feltételek mellett t idő múlva a produkálható részt $a \cdot H(t)$-vel jelöljük. A gátlás folyamatára feltesszük, hogy

$$(5.1) \qquad H(t) = (1-\gamma)^t, \quad \text{ahol} \quad 0 \leqq \gamma < 1.$$

E feltétel szerint tehát egységnyi idő alatt a produkálható anyag γ-szorosa kerül gátlás alá.

A felszabadulásra feltesszük, hogy a tanulástól függetlenül megy végbe, ugyanúgy történik a tanulás ideje alatt mint annak abbahagyása után. Vegyünk fel egy c gátlás alatt levő anyagmennyiséget és tegyük fel, hogy csak a felszabadulás folyamata játszódik le. A még t idő múlva gátlás alatt levő anyagmennyiséget $c \cdot K(t)$-vel jelöljük és feltesszük, hogy

$$(5.2) \qquad K(t) = (1-\delta)^t, \quad \text{ahol} \quad 0 \leqq \delta < 1.$$

δ tehát az egységnyi idő alatt felszabadult anyagmennyiség aránya.

Megjegyezzük, hogy (5.1) és (5.2) nem lényeges feltevések (melyek meglehetősen önkényesek), sőt a tanulás folyamatát ezek figyelembevétele nélkül írjuk le, így csupán egy lehetőségként szerepelnek.

Lássuk be a következő állításokat:

a) Egy a felejtéssel szemben $\varphi(\xi)$ súlyú a produkálható anyagmennyiségből t idő múlva

$$(5.3) \qquad a \cdot \frac{E(\xi+t)}{E(\xi)} \cdot H(t)$$

mennyiség lesz produkálható, ha a felszabadulástól eltekintve csak a felejtést és a gátlást vesszük figyelembe.

b) Jelöljük $a \cdot \overline{K(t) \cdot H(t)}$-val az egy tetszőleges (adott pillanatban) produkálható a anyagmennyiségből t idő elteltével még mindig produkálható anyagmennyiséget, ha a gátlást és a felszabadulást vesszük figyelembe, de a tanuló nem felejt. Igaz az alábbi egyenlőség:

$$(5.4) \qquad a \cdot \overline{K(t) \cdot H(t)} = a \cdot \left[\frac{\ln(1-\delta)}{\ln(1-\delta)(1-\gamma)} + \left(1 - \frac{\ln(1-\delta)}{\ln(1-\delta)(1-\gamma)}\right)((1-\delta)(1-\gamma))^t\right].$$

c) A felejtéssel szemben $\varphi(\xi)$ súlyú c nagyságú gátlás alatt levő anyagmennyiségből t idő múlva

$$(5.5) \qquad c \cdot \frac{E(\xi+t)}{E(\xi)} \cdot K(t)$$

lesz gátlás alatt, ha rá csak a felejtés és a felszabadulás hat.

d) Egy adott pillanatban tanult (egységnyi súlyú) a produkálható anyagmennyiségből t idő múlva

$$(5.6) \qquad a \cdot \overline{H(t) \cdot K(t)} \cdot E(t)$$

lesz produkálható, ha a tanuló felejt, de fellép a gátlás és a felszabadulás is.

A fenti állítások közül csak a b)-t igazoljuk, hisz a többi triviális, illetve d) következménye b)-nek.

A b) állítás igazolásához vegyünk fel egy a produkálható anyagmennyiséget (a felejtéssel szembeni súlya lényegtelen, mivel most a felejtést nem vesszük figyelembe). A gátlásra és a felszabadulásra adott feltevéseink alapján nyilvánvaló, hogy

$$a\cdot\overline{K(t+\Delta t)\cdot H(t+\Delta t)} = a\overline{K(t)\cdot H(t)}\cdot(1-\gamma)^{\Delta t}+a[1-\overline{K(t)\cdot H(t)}]\cdot[1-(1-\delta)^{\Delta t}],$$

mert t-ben $a\cdot\overline{K(t)\cdot H(t)}$ c produkálható és $a\cdot[1-\overline{K(t)\cdot H(t)}]$ a gátlás alatt levő alatt levő anyagmennyiség. Ebből viszont a $\overline{K(t)\cdot H(t)}$ függvényre a

$$(\overline{K(t)\cdot H(t)})' = \overline{K(t)\cdot H(t)}\cdot\ln(1-\delta)(1-\gamma)-\ln(1-\delta)$$

differenciálegyenletet kapjuk, melyet $\overline{K(0)\cdot H(0)}=1$ feltétel mellett megoldva az állításunkat igazoló összefüggéshez jutunk.

Közvetlenül beláthatók az alábbi következmények:

a) A felejtéssel szemben $\varphi(\xi)$ súlyú, c nagyságú gátlás alatt levő anyagmennyiségből t idő alatt

$$c\cdot(1-K(t))\cdot\frac{E(\xi+t)}{E(\xi)} \tag{5.7}$$

válik produkálhatóvá, ha közben a felszabadulás és a felejtés játszódik le gátlás nélkül.

b) Egy adott pillanatban tanult (egységnyi súlyú) a produkálható anyagmennyiségből t idő múlva

$$a\cdot[1-\overline{K(t)\cdot H(t)}]\cdot E(t) \tag{5.8}$$

lesz gátlás alatt, ha feltételezzük a felejtést, gátlást és a felszabadulást is.

Ezek után határozzuk meg az x-ben produkálható $g(x)$ anyagmennyiséget, ha a tanuló $[0, x]$ intervallumon folyamatosan tanul, ahol a felejtés, gátlás és a felszabadulás a fent leírt módon történik.

Mivel a második fejezetben bevezetett $D(x_1, x_2)$ függvénynek most az (5.6) alatti függvény felel meg, ezért $g(x)$-et a következő formában írhatjuk fel:

$$g(x) = \int_0^x \lambda(t)\cdot\overline{K(x-t)\cdot H(x-t)}\cdot E(x-t)\,dt. \tag{5.9}$$

Valamivel bonyolultabb az x-ben produkálható $G(x, y)$ anyagmennyiség felírása, ha a tanuló y ($\leqq x$) ideig tanult folyamatosan, majd abbahagyta a tanulást. A $G(x, y)$ felírása érdekében osszuk a $[0, y]$ intervallumot n részre, az így kapott részintervallumok i-edik tagját jelöljük $[x_i, x_{i+1}]$-el és vizsgáljuk először az ezen tanult anyag változását.

Az $[x_i, x_{i+1}]$ intervallumon tanult $\lambda(\xi)[x_{i+1}-x_i]$ anyagmennyiségből (ahol $x_i\leqq\xi_i\leqq x_{i+1}$) (5.8) alapján y-ban

$$\lambda(\xi_i)\cdot[x_{i+1}-x_i]\cdot[1-\overline{H(y-\xi_i)\cdot K(y-\xi_i)}]\cdot(y-\xi_i)$$

lesz gátlás alatt, mivel a megtanulástól $y-\xi_i$ idő telt el. Az $[y, x]$ intervallumon ez az anyag felejtődik és felszabadul. Mivel y-ban a felejtéssel szemben $\varphi(y-\xi_i)$ súlyú volt, ezért $x-y$ idő alatt

$$\frac{E(x-\xi_i)}{E(y-\xi_i)}\text{-ed részét felejti el,}$$

ugyanakkor felszabadul $K(x-Y)$-ad része, így x-ben már csak

$$\lambda(\xi_i)\cdot[x_{i+1}-x_i]\cdot\left[1-\overline{K(y-\xi_i)\cdot H(y-\xi_i)}\right]\cdot E(x-\xi_i)\cdot K(x-y)$$

mennyiség lesz gátlás alatt.

Az összes részintervallumon tanult anyag x-ben gátlás alatt levő részét összegezve, majd a beosztást minden határon túl finomítva kapjuk az x-ben gátlás alatt levő összes anyagmennyiséget:

$$\int_0^y \lambda(t)\cdot\left[1-\overline{K(y-t)\cdot H(y-t)}\right]\cdot E(x-t)\cdot K(x-y)\,\mathrm{d}t.$$

Ha ezt kivonjuk az x-ben tudott összes $F(x, y)$ anyagból, akkor megkapjuk az x-ben produkálható $G(x, y)$ anyagmennyiséget:

$$(5.10)\qquad G(X, y) = F(x, y)-\int_0^y \lambda(t)\cdot\left[1-\overline{K(y-t)\cdot H(y-t)}\right]\cdot E(x-t)\cdot K(x-y)\,\mathrm{d}t.$$

Vizsgáljuk meg még részletesebben a produkálható anyag változását a tanulás abbahagyása után, azaz rögzítsük y értékét és $G(x, y)$-t x függvényében vizsgáljuk.

A rövidebb írásmód kedvéért vezessük be az

$$1-\overline{K(y-t)\cdot H(y-t)} = q(t).$$

jelölést.

E jelölést felhasználva (5.10)-ből kapjuk, hogy

$$F(x, y)-G(x, y) = \int_0^y \lambda(t)\cdot E(x-t)\cdot q(t)\cdot K(x-y)\,\mathrm{d}t.$$

Ebből az alakból leolvashatjuk a modell egyik fontos tulajdonságát: a tanulás abbahagyása után egy bizonyos idő elteltével a produkálható anyag mennyisége nem különbözik lényegesen a tudott anyagtól, pontosabban

$$(5.11)\qquad \lim_{x\to\infty}[F(x, y)-G(x, y)] = 0,$$

ha y rögzített.

Ennek belátásához megjegyezzük, hogy

$$\int_0^y \lambda(t)\cdot E(x-t)\cdot q(t)\,\mathrm{d}t$$

x-nek korlátos függvénye és $\lim\limits_{x\to\infty} K(x-y) = 0$ (l. (5.2)), ha $\delta\neq 0$.

A modell másik fontos tulajdonsága, hogy a tanulást befolyásoló paraméterek megadhatók úgy, hogy a produkálható anyag mennyisége a tanulás abbahagyása után növekedjék. Ez a növekedés általában csak egy $[y, m]$ intervallumra vonatkozik, ahol m a tanulási folyamat egy jellegzetes pontja, ugyanis itt a legnagyobb a produkálható anyag mennyisége.

Bebizonyítjuk, hogy p megadható úgy, hogy a

$$G(x+\Delta x, y)-G(x, y) > 0$$

teljesüljön rögzített y, x, Δx, γ, δ esetén, ahol feltesszük, hogy $0<\gamma$ és $0<\delta$.

Írjuk fel $G(x, y)$ függvény növekményét (5.10) és (3.2) felhasználásával:

$$G(x+\Delta x, y)-G(x, y) =$$

$$= \int_0^y \lambda(t)\cdot E(x+\Delta x-t)\cdot[1-q(t)K(x+\Delta x-y)]-E(x-t)\cdot[1-q(t)K(x-y)]\}\,dt.$$

Válasszuk p-t akkorára, hogy t egy rögzített t_0 értékére teljesüljön az

$$\left[1-\frac{1}{p\varphi(x-y)}\right]^{\Delta x} > \frac{1-q(t_0)\cdot K(x-y)}{1-q(t_0)\cdot K(x-y)\cdot(1-\delta)^{\Delta x}}$$

egyenlőtlenség (ezt megtehetjük, mivel a bal oldal 1-hez tart, ha p tart a végtelenhez, a jobb oldal értéke pedig 1-nél kisebb). Ekkor viszont minden $0 \leqq t \leqq t_0$ feltételt kielégítő t-re igaz az

$$\frac{E(x+\Delta x-t)}{E(x-t)} > \frac{1q(t)\cdot K(x-y)}{1-q(t)\cdot K(x-y)\cdot(1-\delta)^{\Delta x}}$$

összefüggés, mert a jobb oldali hányadosnál t helyére egyre kisebb számot írva $q(t)$ növekedése miatt a hányados csökken, másrészt az $E(x+\Delta x-t)$ és $E(x-t)$ függvények hányadosa (3.3) alapján csak nagyobb lehet az előző egyenlőtlenség bal oldalánál.

Az utóbbi egyenlőtlenség alapján viszont könnyen belátható, hogy az így választott p esetén a $[0, t_0]$ intervallumon a növekmény integrálja alatti függvény pozitív. Ehhez csak azt kell még észrevenni, hogy $K(x-y)\cdot(1-\delta)^{\Delta x} = K(x+\Delta x-y)$.

Írjuk fel a fenti integrált összeg formájában:

$$G(x+\Delta x, y)-G(x,y) = \int_0^{t_0} + \int_{t_0}^{y},$$

majd t_0-lal közelítsünk y értékéhez, közben p értékét úgy változtatva, hogy az összeg első tagja mindig pozitív maradjon. Mivel az összeg második tagja ilyen feltétel mellett 0-hoz tart, ezért elérhető, hogy az összeg, illetve a növekmény pozitív legyen, ami állításunkat igazolja.

6. EGY KÍSÉRLETI EREDMÉNY ELEMZÉSE

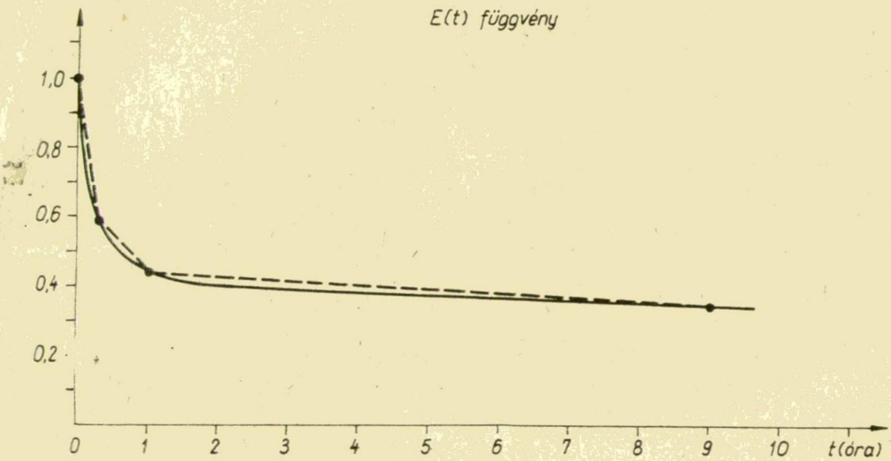
Ebbinghaus a felejtés mértékét vizsgálta (l. Rubinstein: Általános pszichológia alapjai I. kötet) úgy, hogy a kísérleti személyekkel értelmetlen szótagokat tanultatott meg, majd bizonyos időközönként felmérte a még tudott anyagmennyiséget a megtanulás abbahagyása után.

Ebbinghaus kísérleteinek eredményét az alábbi táblázat mutatja, ahol az általunk bevezetett jelöléseket használjuk:

t (óra)	0	$\frac{1}{3}$	1	9	24	48	72	744
$E(t)$	1	0,592	0,442	0,358	0,337	0,278	0,254	0,211

A pszichológiában ezen eredmény realitását vitatják (éppen az értelmetlen tanulás konstruálása miatt), mi ezzel a kérdéssel nem foglalkozunk, a fenti értékeket mint kísérleti eredményeket elfogadjuk és megállapításainkat ez alapján tesszük meg.

A továbbiakban 9 órás folyamatos tanulást tételezünk fel, így az előző táblázat első négy adatára lesz szükségünk. A szóbanlevő négy adat $E(t)$ függvény négy pontját határozza meg, a többi értékeket mi adjuk meg grafikon segítségével, ami



1. ábra

természetesen egy bizonyos mértékig önkényes, amit további kísérlettel lehetne helyesbíteni. $E(t)$ függvényen tehát az 1. ábrán levő grafikon által meghatározott függvényt értjük.

Határozzuk meg először $f(x)$ függvényt $\lambda(t) \equiv 1$ feltétel mellett. $f(x)$-et (3.1) alapján a görbe alatti terület kiszámításával határozzuk meg, melyre az alábbi eredményeket kapjuk:

x (óra)	0	0,1	$\frac{1}{3}$	0,5	$\frac{2}{3}$	1	2	4	9
$f(x)$	0	0,088	0,244	0,343	0,435	0,575	0,995	1,77	3,62

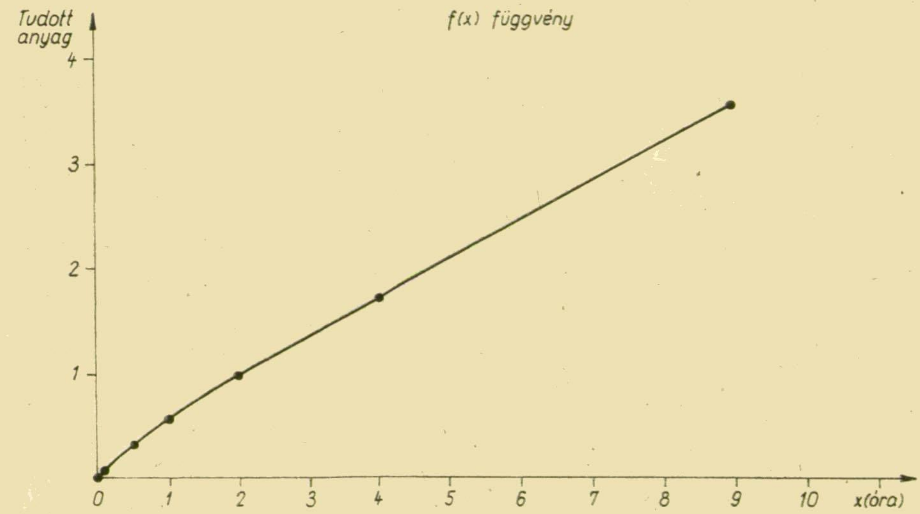
$f(x)$ függvényt a 2. ábrán ábrázoltuk.

Ehhez hasonlóan határozhatnánk meg az $F(x, y)$ függvényt is. Érdekes feladat még a $\varphi(t)$ súlyfüggvény meghatározása. Mivel

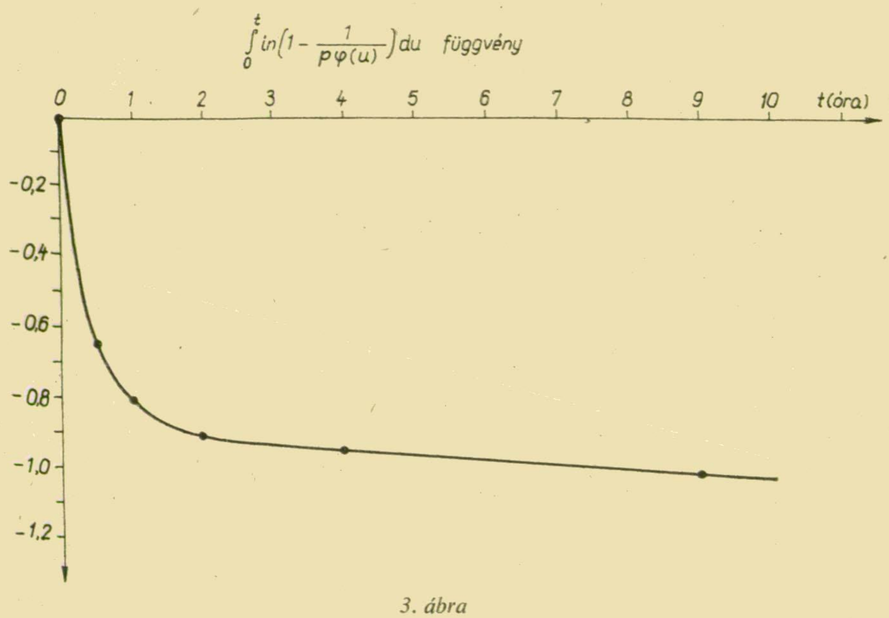
$$E(t) = e^{\int_0^t \ln\left(1-\frac{1}{p\varphi(u)}\right)du},$$

ezért $\varphi(t)$ meghatározását több lépésben végezzük. Először az 1. ábra alapján számítással meghatározzuk a kitevő értékét a fentebb már szerepelt jellegzetes pontokban, majd ezt ábrázoljuk. A részletes számításokat elvégezve az alábbi értéktáblázathoz jutunk:

t (óra)	0	0,1	$\frac{1}{3}$	0,5	$\frac{2}{3}$	1	2	4	9
kitevő	0	−0,274	−0,524	−0,644	−0,702	−0,816	−0,915	−0,959	−1,027



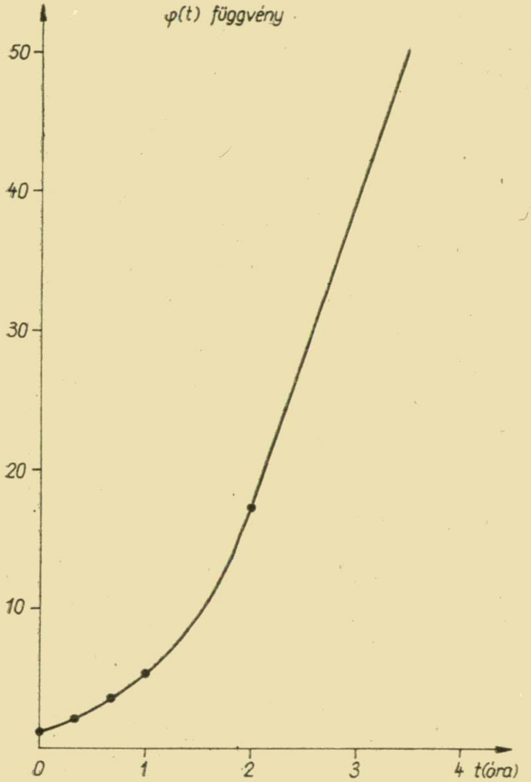
2. ábra



3. ábra

E függvény grafikonja alapján (l. 3. ábra) méréssel meghatározzuk ennek a differenciálhányados függvényét:

t (óra)	0	0,1	$\frac{1}{3}$	0,5	$\frac{2}{3}$	1	2	4	9
$\ln\left(1-\frac{1}{p\varphi(t)}\right)$	-4	$-1{,}92$	$-0{,}66$	$-0{,}46$	$-0{,}35$	$-0{,}22$	$-0{,}06$	$-0{,}02$	–



4. ábra

Mivel $\varphi(0)=1$, ezért $\ln\left(1-\frac{1}{p}\right)=-4$, amiből $p=1{,}02$. A táblázat alapján $\varphi(t)$ értékei az adott pontokban számítással határozhatók meg. A részletes számítások eredményeit az alábbi táblázat tartalmazza, a $\varphi(t)$ függvényt a 4. ábrán ábrázoltuk.

t (óra)	0	0,1	$\frac{1}{3}$	0,5	$\frac{2}{3}$	1	2	4	9
$\varphi(t)$	1	1,195	2,11	2,77	3,46	5,2	17,2	54	–

Megjegyezzük még, hogy az $E(t)$ függvény 0 pont körüli gyors csökkenése az oka annak, hogy p igen közel van az 1-hez.

ТЕОРЕТИКО-ФУНКЦИОНАЛЬНОЕ ОПИСАНИЕ ПРОЦЕССА УЧЕНИЯ

ИШТВАН СЕКЕРЕШ

ANALYTIC DESCRIPTION OF THE PROCESS OF LEARNING

I. SZEKERES

KÜLÖNLENYOMAT

Magyar Pszichológiai Szemle

A MAGYAR TUDOMÁNYOS AKADÉMIA
PSZICHOLÓGIAI BIZOTTSÁGA
ÉS
A MAGYAR PSZICHOLÓGIAI
TÁRSASÁG FOLYÓIRATA

SZEKERES ISTVÁN

EGY SPECIÁLIS TANULÁSI FOLYAMAT VIZSGÁLATA INFORMÁCIÓELMÉLETI MÓDSZEREKKEL

1976 XXXIII. KÖTET 5. SZÁM

SZEMLE

EGY SPECIÁLIS TANULÁSI FOLYAMAT VIZSGÁLATA INFORMÁCIÓELMÉLETI MÓDSZEREKKEL

SZEKERES ISTVÁN
Toldi Miklós Élelmiszeripari Szakközépiskola Nagykőrös

Napjainkban sok szó esik a pszichológia és a matematika kapcsolatáról, illetve a matematikai módszerek pszichológián belüli alkalmazásáról. Az alkalmazhatóság nem vita tárgya, de az alkalmazás köre és mértéke még messze elmarad a lehetőségek teljes kihasználásától. Az azonban már látható, hogy a pszichológia és a matematika említett kapcsolatában (felépítése következtében) kiemelt szerepet kell tulajdonítanunk az információelméletnek. Bizonyos feltételek mellett ugyanis egy adott inger matematikailag leírható információként definiálható. Hasonló kapcsolat létesíthető a reakció és az entrópia között olyan értelemben, hogy a reakció bizonytalanságát a matematikában definiált entrópiával mérhetjük. Különösen hasznosnak bizonyul ez a felfogás a tanulási folyamatok leírásánál, ahol kézenfekvő a tanulást mint entrópiacsökkenéssel (információmennyiség növekedésével) járó folyamatot tekinteni.

Fentiek is indokolják azt a tényt, hogy a pszichológiában egyre gyakrabban szerepel az entrópia és a matematikai értelemben vett információmennyiség fogalma. E matematikai fogalmak pszichológián belüli alkalmazása azonban igen körültekintő munkát igényel. Egyik alapvető probléma, hogy egy adott ingerhez nem rendelhetünk egyértelműen információmennyiséget, mert nem lehet kiküszöbölni az egyén szerepét, annak speciális tulajdonságait. Például egy gép valamely alkatrészének megtekintése elegendő információt adhat a szakember számára a gép megjavításához, ugyanakkor ez az inger a laikus számára semmiféle információt nem szolgáltat. Hasonló probléma merül fel az entrópiával kapcsolatban is, hisz (matematikailag) ugyanaz a valószínűségeloszlás két különböző személy esetén nem biztos, hogy ugyanazt a (pszichológiai értelemben vett) bizonytalanságot hozza létre. Az is előfordulhat, hogy két különböző formában megjelenő, de matematikailag megegyező valószínűségeloszlás ugyanazon személy esetén is különböző bizonytalanságot hoz létre.

Az említett problémák megoldása, tehát a matematikai és pszichikai entrópia, illetve információmennyiség kapcsolatának feltárása után csakis kísérletek alapján történhet. E dolgozat célja egy ilyen témájú kísérlet elemzése, melynek alapgondolatát C. Günter munkája adta. (Megjelent a Pädagogik 1965/20 számában, magyar nyelvű fordítása a Tanulmányok a programozott tanítás köréből című egyetemi jegyzetben található.)

C. Günter dolgozatában leír egy sakktábla-kísérletet, mely szerint a kísérletvezető a sakktáblán található 64 mező valamelyikén elhelyez egy tárgyat, majd a kísérleti alanynak ezt alternatív kérdések alapján kell megtalálnia. A szerző egyrészt a kérdések „jóságát" vizsgálja és (információelméleti fogalmak felhasználásával) arra az eredményre jut, hogy a kísérleti alanynak akkor a legjobb a taktikája, ha a keresési területet mindig felezi, azaz ha a keresési terület valamely felére kérdez. (Ezzel a problémával e cikkben is részletesen foglalkozunk, C. Günter dolgozatának áttanulmányozása azonban nagyban hozzásegíti az olvasót az információelméleti meggondolásaink könynyebb megértéséhez is.) A kísérlet előnye, hogy a taktika jóságát igen könnyű matematikailag leírni, ami viszont e fogalom mérését teszi lehetővé. C. Günter felveti továbbá azt a kérdést, hogy egy kísérleti alany taktikája milyen kapcsolatba hozható annak más pszichikai tulajdonságával. A felvetett pszichológiai témák közül itt a tanulási folyamat problémáját emeljük ki. E kísérlet folyamán ugyanis a kísérleti alany „megtanulja" az elhelyezett tárgy helyét, így a taktika jóságával e speciális tanulási folyamat minőségét mérjük. E kísérlet egyúttal igen szépen példázza azt a már említett tényt, hogy a tanulás egy entrópiacsökkenéssel járó folyamat. Egyszerűsége azonban nagyon szűk határt szab alkalmazhatóságának és a belőle levonható következtetéseknek, ezért tűztük célul e kísérlet általánosítását. Az alább ismertetésre kerülő kísérlet alkalmazási köre jóval túlhaladja az előbbiét, tehát bővebb azoknak a tanulási folyamatoknak a halmaza melyeket a következő kísérlet alapján tanulmányozhatunk.

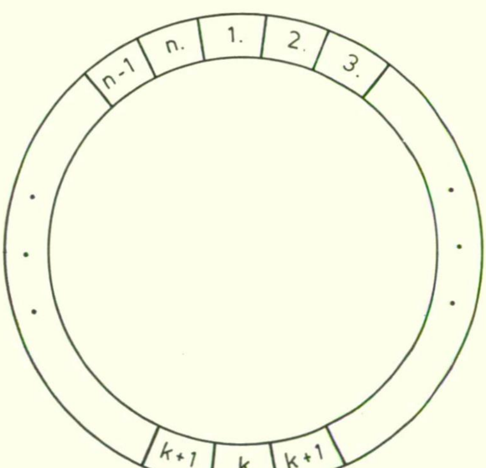
Hogy a nem matematikus olvasók számára könnyebben érthető legyen mondanivalónk, az általánosítást két lépésben végezzük. Első lépésben az illető tárgy elhelyezése lehetőségeinek számát választjuk (64 helyett) tetszőlegesnek, az egyes lehetőségek valószínűségeit viszont továbbra is egyenlőnek tételezzük fel. A második lépésben már az elhelyezési lehetőségek (elemi események) valószínűségeloszlását is tetszőlegesnek választjuk.

A. C. Günter-féle sakktábla kísérlet általánosítása

A) Egyenlő valószínűségek esete

Az említett dolgozatban leírt sakktábla-kísérletet először egy vele lényegében egyenértékű kísérlettel helyettesítjük.

Vegyünk fel egy körgyűrűt és osszuk ezt n egyenlő részre (1. ábra), így n „rekeszt" kapunk, melyeket 1., 2., . . . ,n. sorszámokkal látunk el. (A rekeszek ilyen elrendezésének előnye a sakktáblával szemben egyrészt az, hogy itt geometriailag sincs különbség az egyes rekeszek között, másrészt ennek könnyebb lesz az általánosítása.)



1. ábra

Végezzük el a következő kísérletet: Valamilyen módon jegyezzünk meg egy rekeszt (például a k-adikat), majd kerestessük meg ezt egy kísérleti alannyal a következő feltételek mellett:

a) A kísérleti alany csak alternatív kérdést tehet fel a válasz tehát csak „igen" (I), vagy „nem" (N) lehet.

b) A kísérleti alany nem találgathat. Ez azt jelenti, hogy csak akkor mondja a végső választ (akkor ér véget a kísérlet), ha biztosan tudja, hogy melyik a megjelölt rekeszt. Ha például a megtalálás jutalommal jár, akkor ezt úgy kell megválasztani, hogy e feltétel biztosítva legyen, azaz a biztos válaszra ösztönözzön.

c) A rekeszek megjelölésének valószínűsége a kísérleti alany előtt legyen egyenlő. A kísérleti alanyt ne befolyásolja olyan tényező, mely számára egyes rekeszek esélyeit növelné, vagy csökkentené. Ezt például úgy érhetjük el, hogy a számot húzzuk.

d) A kísérleti alany arra törekedjen, hogy minél kevesebb kérdés után találja meg a megjelölt rekeszt. Az időt nem figyeljük.

Mivel a kísérleti alany a rekeszek valamely részhalmazára kérdezhet, ezért az összes lehetséges kérdések száma 2^n. E kérdések azonban a kérdés jósága szempontjából nem mind különbözőek. Könnyű ugyanis belátni, hogy ugyanazt tudja meg az alany akár az előző 10-re, akár az utolsó n—10-re kérdez. A kérdés jósága szempontjából nincs különbség két kérdés között akkor sem, ha az egyik az első kettő, a másik pedig az utolsó kettő rekeszre kérdez, mert (mint látni fogjuk) a várható entrópiacsökkenés mindkét esetben ugyanannyi. Ahhoz, hogy e problémákat pontosabban megfogalmazhassuk és vizsgálat tárgyává tehessük mindenekelőtt algebrailag kell leírni a kísérletet.

Tekintsük elemi eseménynek, ha éppen a k-adik rekesz lett megjelölve és jelöljük ezt w_k-val, így n elemi eseményt definiáltunk, ezek halmazát W-vel jelöljük. Vezessük be a szokásos műveleteket (e dolgozat megértéséhez szükséges valószínűségszámítási ismeretek megtalálhatók Rényi Alfréd Valószínűségszámítás c. tankönyvében 15.—106. old.), majd W-t bővítsük Boole-algebrára, amit eseményalgebrának tekintünk. Az így kapott eseményalgebra általános elemét q-val jelöljük, az összes elemek száma 2^n. Mivel ez az eseményalgebra izomorf a rekeszek részhalmazai által alkotott halmazzal, így a lehetséges kérdések halmazával is, nem okoz félreértést, ha a q elemeket kérdéseknek is nevezzük, vagyis ha a kísérleti alany a q eseményre kérdez, akkor azt mondjuk, hogy a q kérdést teszi fel.

A fent definiált eseményalgebrán egy valószínűségi mezőt értelmezünk, melyre c.) feltétel alapján a következő kikötést tesszük:

$$P(w_k) = \frac{1}{n} \quad \text{minden k-ra.}$$

Jelöljük továbbá I/q-val azt az eseményt, hogy a kísérleti alany a q kérdést teszi fel (q eseményre kérdez) és I választ (megerősítést) kap. Így egy újabb Q-val jelölt eseményalgebrához jutottunk, melynek elemi eseményei az

$$I/w_1,\ I/w_2, \ldots ,\ I/w_n \text{ események.}$$

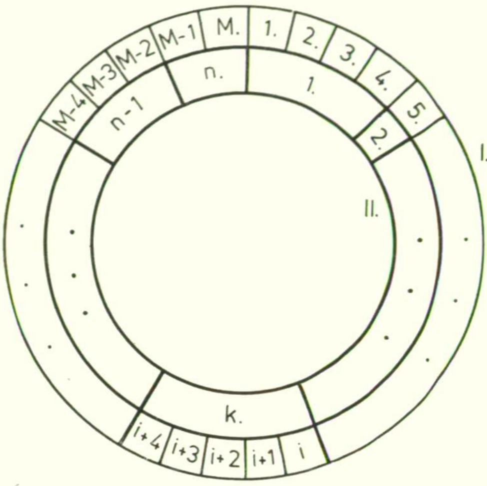
Feltesszük, hogy $P(I/q) = P(q)$ és ezzel a $P\times Q$ valószínűségi mezőt definiáltuk.

$\overline{I/q}$ jelölés helyett az N/q jelölést használjuk, mely azt a nyilvánvaló tényt fejezi ki, hogy ha az alany q kérdés feltevésekor nem az I megerősítést kapja, akkor N-t.

A $P\times Q$ valószínűségi mező entrópiáját $H(Q)$-val jelöljük és ezt a kísérleti alany kezdeti bizonytalanságának nevezzük. E definíció alapján a kiszámíthatjuk a kísérleti alany kezdeti bizonytalanságát:

Mindenekelőtt bemutatunk néhány példát ennek gyakorlati megvalósítására:

1. Vegyünk fel most is egy körgyűrűt és osszuk ezt M egyenlő részre, majd a felvett körgyűrűvel koncentrikusan egy újabb körgyűrűt helyezünk el és ezt n(W $\leq$ M) részre osztjuk oly módon, hogy az eredeti gyűrű bizonyos beosztásait meghosszabbítjuk (2. ábra). Az új körgyűrű (II.) minden beosztásához az eredeti (I.) néhány (de legalább egy) beosztása tartozik. A II.-vel jelölt körgyűrű k-adik beosztásához tartozó I.-beli beosztások számát i_k-val jelöljük, ezek összege M. Véletlenszerűen (például a szám húzásával) jelöljük meg az I.-es gyűrű valamely rekeszét. Most azonban nem a konkrétan megjelölt I.-beli rekeszt kell a kísérleti alanynak megtalálnia, hanem csak azt kell eldöntenie, hogy a megjelölt I.-beli rekesz a II.-es gyűrű melyik sávjába esik. A megjelölés tehát az I.-es gyűrű alapján történik, de az alanynak a II.-es gyűrű sávjaira kell kérdeznie.



2. ábra

Annak valószínűsége, hogy a megjelölt I.-beli rekesz éppen a II.-es gyűrűk k-adik sávjába esik.

$$p_k = \frac{i_k}{M}.$$

E konstrukció előnye, hogy igen szemléletes és tetszés szerint variálhatók az egyes rekeszek megjelölésének valószínűségei mind sorrendben, mind nagyságban.

2. Ha a körgyűrűnél akarunk maradni, akkor az egyes rekeszek különböző valószínűségű megjelölését úgy is elérhetjük, hogy csupán egy körgyűrűt alkalmazunk és ezt (mint az egyenlő valószínűségű esetben) n egyenlő részre osztjuk, de azoknak a számoknak a halmazát melyekből a megjelölendő rekesz sorszámát húzzuk megváltoztatjuk. Az urnába (amiből a számot húzzuk) most nem egy-egy sorszámot helyezünk el, hanem az 1.-es sorszámból i_1-et, a 2.-es sorszámból i_2-t, stb. Azt, hogy az egyes sorszámokból mennyi van az urnában a kísérleti alany tudomására kell hozni.

Annak valószínűsége, hogy a k-adik rekesz lesz megjelölve

$$p_k = \frac{i_k}{M}, \text{ ahol } M = \sum_{k=1}^{n} i_k.$$

Nyilvánvaló, hogy matematikai szempontból a kísérlet ilyen formában való végrehajtása ekvivalens az előbbivel, pszichológiai szempontból mégis különbözik attól, mert itt geometriailag nincs szemléltetve az egyes rekeszek megjelölésének valószínűsége.

3. Végül még egy, az előbbiektől formailag különböző kivitelezési lehetőséget mutatunk be. Vegyünk n fiókot és M számú golyót. A M golyó közül egy piros, a többi fehér legyen. A golyók színének figyelése nélkül rakjunk a fiókokba rendre $i_1, i_2, \ldots, i_n$ számú golyót. A kísérleti alany tájékoztatása érdekében ezeket a számokat a megfelelő fiókokon feltüntetjük.

A kísérleti alany feladata a piros golyó helyének megtalálása. A kísérletvezető természetesen már a kérdések megtétele előtt ismeri a piros golyó helyét.

Annak valószínűsége, hogy a piros golyó a k-adik fiókba került most is

$p_k = \frac{i_k}{M}$ lesz, ahol hallgatólagosan feltettük, hogy az összes M számú golyót elhelyeztük a fiókba.

Ezek után algebrailag elemezzük a kísérlet ilyen feltételek melletti lehetséges kimeneteleit. Lényegében most is azt vizsgáljuk, hogy az egyes kérdések után hogyan változik az entrópia értéke.

A $P \times Q$ valószínűségi mező entrópiája (a kezdeti entrópia) ebben az esetben

$$H(Q) = \sum_{k=1}^{n} p_k \cdot \log_2 \frac{1}{p_k}$$

lesz, mely kisebb, vagy egyenlő mint $\log_2 n$. Ez a a kérdések szempontjából azt jelenti, hogy jó taktikával „átlagosan" kevesebb kérdés után lehet megtalálni a megjelölt rekeszt, ha az egyes rekeszek megjelölésének valószínűségei különbözőek.

Tegyük fel, hogy a kísérleti alany egy olyan q kérdést tesz fel, melynek valószínűsége u, tehát

$P(I/q) = u$, ahol $u = p_{i1} + p_{i2} + \ldots + p_{ir}$ (r számú rekeszre kérdezett)

Ha a kísérleti alany a q kérdés feltevése után I megerősítést kap (ennek valószínűsége u), akkor az új Q' eseményalgebrát az r számú rekesz generálja, amelyekre az alany rákérdezett. Feltesszük, hogy az így kapott eseményalgebra eseményeinek valószínűségei rendre

$$\frac{p_{i1}}{u}, \frac{p_{i2}}{u}, \ldots, \frac{p_{ir}}{u}$$

lesznek, tehát az új eloszlás az r rekesz eredeti valószínűségeivel arányos. Ilyen feltétel mellett a fennmaradó entrópia értéke

$$\sum_{k=1}^{r} \frac{1}{u} \cdot p_{ik} \cdot \log_2 u \cdot \frac{1}{p_{ik}}$$

Ehhez hasonlóan járhatunk el abban az esetben is, ha az u valószínűségű kérdés feltevése után N megerősítést kap a kísérleti alany (ennek valószínűsége $1-u$). Ekkor megint egy újabb, de az előbbitől különböző eseményalgebrához jutunk, ahol az elemi események száma $n-r$, ezek valószínűségei rendre

$$\frac{p_{j1}}{1-u}, \frac{p_{j2}}{1-u}, \ldots, \frac{p_{j(n-r)}}{1-u},$$

egyenletet, amelyből ΔH_u értékének kifejtésével a következő összefüggést kapjuk S-re:

$$S = \frac{H(Q)}{\log_2 \frac{1}{u^u \cdot (1-u)^{(1-u)}}}.$$

Ebből leolvashatjuk, hogy a kérdések várható száma akkor lesz a minimális, ha $u = \frac{1}{2}$ és a minimum értéke $H(Q)$. Jó taktika esetén tehát a kérdések várható száma éppen a kezdeti entrópiával egyezik meg.

A kérdések várható száma és a tényleges kérdésszám között hasonló a kapcsolat, mint a várható entrópiacsökkenés és a bekövetkezett entrópiacsökkenés között. Sokkal bonyolultabb lenne a kérdés, ha nem tennénk azt (az egyébként önkényes) a megszorítást, hogy a kísérleti alany „mindig u valószínűségű kérdést tesz fel" Nálunk ez a megszorítás nem fog problémát okozni, mert a fenti összefüggést arra akarjuk felhasználni, hogy segítségével megállapítsuk, a kísérleti alany „átlagosan" milyen valószínűségű kérdéseket tett fel. A kezdeti entrópia és a tényleges kérdésszám hányadosából ugyanis a fenti összefüggés alapján u értéke meghatározható. Az így kapott u értéket nevezhetjük a kísérleti alany által feltett kérdések átlagos valószínűségének. Mivel itt várható értékekkel dolgoztunk, ezért ez csak akkor ad reális képet, ha S helyére nagyszámú kísérlet során bekövetkezett kérdésszámok átlagát helyettesítjük, vagy ha a kezdeti entrópia elég nagy, tehát ha nagyszámú kérdés után lehet megtalálni a megjelölt rekeszt.

A kísérlet gyakorlati végrehajtásáról

Más kísérletekhez hasonlóan a szóban levő kísérlet is három fő lépésből tevődik össze: a kísérlet megtervezése, végrehajtása és értékelése.

Vizsgáljuk meg e lépéseket kissé részletesebben.

1. A kísérlet megtervezése
   a) A kísérlet formájának megtervezése
   - Hány elemi eseményt és mekkora valószínűségértékekkel definiáljunk?
   - A tervezett valószínűségi mezőt milyen formában „tálaljuk" a kísérleti alany elé?

   (E kérdések pszichológiai szempontból nagyon fontosak.)

   b) Előzetes matematikai elemzés
   - Az elemi események valószínűségeinek meghatározása.
   - A kezdeti entrópia kiszámítása.
   - A legjobb taktikai lehetőségek áttekintése.
2. A kísérlet végrehajtása
   a) A kísérleti alannyal ismertetni kell a feladatát
   (Nagyon fontos kihangsúlyozni, illetve megértetni, hogy az elemi események bármely részhalmazára kérdezhet.)

   b) A kísérlet folyamata
   - A kísérletvezető véletlenszerűen megjelöl egy rekeszt.
   - A kísérleti alany alternatív kérdéseket tesz fel.
   - A kísérletvezető sorrendben rögzíti a feltett kérdéseket, vagyis azt, hogy a rekeszek mely részhalmazaira kérdezett a kísérleti alany.

3. A kísérlet értékelése

A) Matematikai értékelés.

a) Az egyes kérdések valószínűségeinek meghatározása.

b) Az egyes kérdések jóságának kiszámítása.

c) Ha szükséges, a kérdések átlagos valószínűségének kiszámítása.

B) Pszichológiai elemzés

A matematikai eredmények alapján, a célnak megfelelően.

A fentiek illusztrálására leírunk egy ténylegesen végrehajtott kísérletet, anélkül, hogy ennek kimeneteléből bármiféle következtetést kívánnánk levonni. (Ehhez természetesen egy, vagy néhány kísérlet elvégzése nem is lenne elegendő.)

A kísérlet alanya egy negyedik osztályos közepes képességű szakközépiskolai tanuló volt.

A kísérlethez 10 darab üres gyufásdobozt és 100 gyufaszálat készítettünk elő. A gyufaszálak közül 99 barna, egy pedig piros színű volt. A dobozokra 1-től 10-ig sorzámokat írtunk, másrészt fetüntettük, hogy az egyes dobozokba hány szál gyufát fogunk elhelyezni. Ezek rendre a következők voltak:

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
5	25	1	2	10	40	1	4	10	2

A kísérleti alany feladata (a gyufaszálak véletlenszerű, a fenti számoknak megfelelő elhelyezése után) a piros gyufaszál helyének megtalálása.

Az így definiált tíz elemi esemény valószínűségei:

1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
0,05	0,25	0,01	0,02	0,1	0,4	0,01	0,04	0,1	0,02

Ez alapján a kezdeti entrópia értéke

$$H(Q) = 0{,}05 \cdot \log_2 \frac{1}{0{,}05} + 0{,}25 \cdot \log_2 \frac{1}{0{,}25} + 2 \cdot 0{,}01 \cdot \log_2 \frac{1}{0{,}01} +$$
$$+ 2 \cdot 0{,}02 \cdot \log_2 \frac{1}{0{,}02} + 2 \cdot 0{,}1 \cdot \log_2 \frac{1}{0{,}1} + 0{,}4 \cdot \log_2 \frac{1}{0{,}4} + 0{,}04 \cdot \log_2 \frac{1}{0{,}04}.$$

Mivel $\log_2 x = \frac{\lg x}{\lg 2}$ (ahol lg 10-es alapú logarimust jelöl), ezért $H(Q)$ értékét a következőképpen számíthatjuk ki:

$$H(Q) = \frac{1}{\lg 2} \cdot [0{,}05 \cdot \lg 20 + 0{,}25 \cdot \lg 4 + 0{,}02 \cdot \lg 100 + 0{,}04 \cdot \lg 50 +$$
$$+ 0{,}2 \cdot \lg 10 + 0{,}4 \cdot \lg 2{,}5 + 0{,}04 \cdot \lg 25].$$

A megfelelő logaritmusokat táblázatból kikeresve és a kijelölt műveleteket elvégezve kapjuk, hogy

$$H(Q) = 2{,}45,$$

tehát jó taktika esetén átlagosan 2—3 kérdés után lehet megtalálni a megfelelő dobozt.

A valószínűségeloszlásból azt is megállapíthatjuk, hogy első kérdésnek valamely 0,5 valószínűségű kérdés a legjobb (pl. ha az 5., 6., vagy 1., 2., 5., 9. dobozra kérdez), tehát az első kérdésnél

$$\Delta H_{max} = 1.$$

A kísérlet vezetője a megfelelő számú gyufaszálat úgy helyezi el a dobozokba, hogy sem ő, sem a kísérleti alany nem látja annak színét. Ezek után a kísérlet vezetője megkeresi a piros gyufaszál helyét (ez a 9. doboz volt) és a dobozokat kör alakban helyezi el.

A kísérleti alany kérdései és a válaszok:

Első kérdés: az 1., 2., 3., 4. dobozra kérdez az alany

Válasz: nem

Második kérdés: az 5., 6., 7., 8. dobozra kérdez az alany

Válasz: nem

Harmadik kérdés: a 9. dobozra kérdez az alany

Válasz: igen

Matematikai elemzés:

Az első kérdés valószínűsége 0,05 + 0,25 + 0,01 + 0,02 = 0,33, az ezután várható entrópiacsökkenés

$$\Delta H_{0,33} = \log_2 \frac{1}{0{,}33^{0,33} \cdot 0{,}67^{0,67}} = 0{,}91, \quad \text{a kérdés jósága } J = 0{,}91.$$

Mivel az első kérdésre tagadó választ kapott az alany, a keresési terület az 5., 6., 7., 8., 9. és 10. dobozokra korlátozódott, azaz a fennmaradt eseményalgebra hat elemi eseményből áll, melyek valószínűségeit ki kell számítani. Annak valószínűsége például, hogy a piros gyufaszál az 5. dobozban helyezkedik el $\frac{10}{67}$, mert ebben 10 szál gyufa van és a még szóba jöhető gyufák száma 67. Hasonlóan számíthatjuk ki a többi elemi esemény valószínűségét is. Ezekre a következő eredményeket kapjuk:

5.	6.	7.	8.	9.	10.
0,15	0,595	0,015	0,06	0,15	0,03

A most feltehető legjobb kérdés valószínűsége 0,595, tehát

$$\Delta H_{max} = \Delta H_{0,595} = 0{,}97$$

Azzal, hogy a kísérleti alany az 5., 6., 7., 8. dobozokra kérdezett 0,82 valószínűségű kérdést tett fel, ezért a várható entrópiacsökkenés

$$\Delta H_{0,82} = 0{,}68, \quad \text{a kérdés jósága pedig}$$

$$J = \frac{0{,}68}{0{,}97} = 0{,}7.$$

Mivel a második kérdés után már csak két elemi eseményből álló eseményalgebra maradt fenn, ezért a harmadik kérdés vizsgálatának nincs különösebb értelme.

E kísérlettel kapcsolatban megjegyezzük, hogy a kísérleti alanyt észrevehetően befolyásolja a dobozok sorszámozása a kérdések megtételében, ajánlatosabb lett volna tehát a dobozok megjelölésére más módot találni.

Pszichológiai vonatkozású megjegyzések a kísérlethez

Az általunk leírt kísérlet pszichológiájának részletes vizsgálata, módszeresen végrehajtott kísérletek eredményeinek elemzése egy önálló dolgozat tárgya lehetne, ezért itt csak utalni tudunk néhány problémára.

Már említettük, hogy a szóban levő kísérletet, mint speciális tanulási folyamatot tekintjük. Láttuk, hogy e folyamat matematikai eszközökkel jól jellemezhető, aminek alapja az, hogy megadható egy ideális eljárás, amihez a tényleges eljárást hasonlítani tudjuk. Az általunk bevezetett matematikai apparátust azonban csak mint segédeszközt használhatja a pszichológus, e folyamat pszichológiai elemzése még számtalan további problémát vet fel.

Pszichológiai szempontból már a kísérlethez szükséges eseményalgebra formai kivitelezése is fontos kérdés. Az általánosított kísérlet matematikai elemzése előtt bemutattunk néhány példát ennek gyakorlati megvalósítására, a kísérletező pszichológus azonban sok, ezektől formailag lényegesen különböző (de matematikailag ezekkel egyenértékű) kísérlet elemzésére használhatja az elmondottakat. C. Günter például bemutat egy beszédmentes kísérletet, amely matematikailag ugyanígy elemezhető.

A kísérlet formai kivitelezése (a kísérleti alany intelligenciaszintjétől függően) lényegesen befolyásolhatja a kérdések megtételét. A sorba szedett rekeszek például sugallhatják, hogy a kísérleti alany ilyen sorrendben kérdezzen. Másrészt a formai kivitelezéstől függ az is, hogy a kísérleti alanyban milyen „szubjektív valószínűségeloszlás" jön létre. Nem mindegy például, hogy ha csak a rekeszekre írt számokkal utalunk azok megjelölésének valószínűségére, vagy ha ezt geometriailag is érzékeltetjük.

A kísérlet formáját tehát a kísérlet célja határozza meg, de figyelembe kell venni a kísérleti alany adottságait is.

A kísérlet természetéből folyó (a bevezetőben is említett) kérdés, hogy egy kísérleti alany taktikájának jósága milyen kapcsolatba hozható annak más pszichikai tulajdonságával?

E kérdés megválaszolása megfelelő számú kísérlet elvégzése után nyert statisztikai adatok alapján nem látszik problémásnak. Ha sikerülne ilyen kapcsolatokat feltárni, akkor e kísérlet egyszerű, metematikailag is elemezhető módszert adna bizonyos pszichikai tulajdonságok vizsgálatához. C. Günter többek között a matematikai készség és a kérdés jóságának kapcsolatát említi. Azt természetesen nem várhatjuk, hogy a kérdés jóságából meghatározhassuk egy személy matematikai készségét, de azt igen, hogy annak egy komponenséről ez alapján hű képet alkothassunk.

Valószínűnek látszik, hogy azonos eseményalgebrával végzett kísérletek egymásutánjakor a kérdések jósága pozitív irányban fog változni, mert a rossz kérdések után átlagosan több lépés szükséges a megjelölt rekesz megtalálásához, amit a kísérleti alanynak idővel fel kell ismernie. Ha ez kísérletileg valóban kimutatható, akkor a kérdések jóságának változása egy újabb tanulási folyamatként fogható fel.

A kísérlet még további, az előbbieknél jóval bonyolultabb tanulási folyamat vizsgálatára is ad lehetőséget. Ha ugyanis a kísérleti alannyal csak az elemi eseményeket ismertetjük, de azok valószínűségeloszlását nem, akkor a kísérlet többszöri elvégzésével a kísérleti alany elé tűzhetjük azt a feladatot, hogy „tanulja" meg az elemi események valószínűségeit, amit természetesen a kísérletvezető nem változtat meg a kísérlet folyamán. Sőt a fenti gondolatmenetet folytatva a kísérletet úgy is általánosíthatjuk, hogy a valószínűségeloszlást (melyet nem közlünk az alannyal) valamilyen törvényszerűség szerint változtatjuk, ekkor a kísérleti alanynak az előbbieken kívül ezt a törvényszerűséget is fel kell ismernie. A kísérlet ilyen formában való végrehajtása nagyon bonyolultnak látszik (és valóban bonyolulttá tehető), azonban bizonyos egyszerűsítésekkel elérhetjük, hogy végrehajtható legyen. Gondoljunk például arra az esetre, amikor a valószínűségeloszlás sorozatot úgy választjuk meg, hogy egy elemi esemény valószínűsége mindig 1, a többié pedig 0, azaz ha előre megtervezzük, hogy melyik rekeszt fogjuk megjelölni.

KÜLÖNLENYOMAT

Magyar Pszichológiai Szemle

A MAGYAR TUDOMÁNYOS AKADÉMIA
PSZICHOLÓGIAI BIZOTTSÁGA
ÉS
A MAGYAR PSZICHOLÓGIAI
TÁRSASÁG FOLYÓIRATA

SZEKERES ISTVÁN

SZUBJEKTÍV VALÓSZÍNŰSÉGELOSZLÁS KÜLÖNBÖZŐ TÍPUSÚ FELADATHELYZETEKBEN

1979 XXXVI. KÖTET 3. SZÁM

SZUBJEKTÍV VALÓSZÍNŰSÉGELOSZLÁS KÜLÖNBÖZŐ TÍPUSÚ FELADATHELYZETEKBEN

SZEKERES ISTVÁN
Toldi Miklós Élelmiszeripari Szakközépiskola, Nagykörös

Előfordul, hogy valamely kérdésre nem tudunk egyértelmű választ adni, hanem több választ is lehetségesnek tartunk. A lehetségesnek tartott $A_1, A_2, \ldots, A_n$ válaszokat teljes szubjektív eseményrendszernek nevezzük. Ehhez definiálhatunk szubjektív valószínűség-eloszlást úgy, hogy az eleget tegyen a valószínűségeloszlással szemben támasztott matematikai feltételeknek. Az így definiált valószínűségek az egyes válaszokra ösztönző ingerek erősségével arányosak. Alapfeladat a szubjektív valószínűségeloszlás felmérése és a mérési eredmények kiértékelése, amit a dolgozat több konkrét példa, elvégzett kísérlet alapján mutat be. Az elvégzett kísérletek fő célja a szubjektív valószínűségeloszlás létezésének alátámasztása és a szubjektív valószínűség stabilitásának vizsgálata. A kísérletek elvégzése, illetve kiértékelése során felmerül néhány további pszichológiai probléma. Ezek közül egyik legérdekesebb az, hogy az eseményrendszer formai oldala mennyiben befolyásolja a szubjektív valószínűségeloszlást.

Pszichológiai kísérleteknél vagy pedagógiai felméréseknél előfordulhat, hogy a kísérleti személy egy adott kérdésre nem tud egyértelmű választ adni, több választ is lehetségesnek tart. Ha ennek ellenére konkrét válasz adására késztetjük, akkor természetesen azt a választ adja, ami szerinte a „legvalószínűbb". Így azonban a kísérletvezető még nem látja világosan, hogy a kisérleti személy miként vélekedik az adott kérdéssel kapcsolatban. A felmérés akkor lenne teljes, ha feltérképeznénk azt, hogy az adott kérdésre mely válaszokat tartja lehetségesnek a kisérleti személy és ezek mekkora „valószínűséggel" kerülnének válaszadásra.

Tegyük fel, hogy a kísérleti személy a feltett kérdésre az $A_1, A_2, \ldots, A_n$ válaszokat és csak ezeket tartja lehetségesnek. Tegyük fel továbbá, hogy ezek a válaszok teljes eseményrendszert alkotnak, azaz ezek közül egy és csak egy a helyes válasz. (Ha a kisérleti személy a valóban helyes választ nem tartja lehetségesnek, azt a kiértékeléskor akkor is az A_i halmazba kell sorolnunk, és akkor azt mondjuk, hogy a helyes válasz 0 valószínűséggel kerül válaszadásra.)

A fentiek alapján a kisérleti személy által lehetségesnek tartott válaszokat egy matematikai értelemben vett teljes eseményrendszerként kezelhetjük, mivel azonban ennek felépülése függ a kisérleti személytől, ezért szubjektív eseményrendszernek nevezzük.

Alapgondolatunk, hogy az egyes válaszokra ösztönző belső kényszer erőssége számokkal jellemezhető. Mivel relatív mérésről van szó, ezért csak a számok aránya a mérvadó. Ha e számokat úgy adjuk meg (a végtelen sok lehetőség közül egyet kiválasztva), hogy értékük 0-nál ne legyen kisebb és 1-nél ne legyen nagyobb, összegük pedig 1-et adjon, akkor e számokat az egyes válaszok szubjektív valószínűségének, e számok halmazát pedig szubjektív valószínűségeloszlásnak nevezzük. Az egyes válaszok szubjektív valószínűségeit q_i-vel fogjuk jelölni ($i = 1, 2, \ldots, n$).

A szubjektív valószínűség az irodalomban többszörösen leírt és vizsgált fogalom, mi inkább a szubjektív valószínűségeloszlás vizsgálatára irányítjuk figyelmünket, mivel így általánosabban vethetők fel bizonyos pszichológiai problémák.

Célunk a szubjektív valószínűségeloszlás felmérésének, elemzésének leírása és bemutatása. Ennek érdekében végeztünk el néhány kísérletet. Mivel e kísérletek inkább illusztrációként szerepelnek, nem vonhatunk le belőlük mélyreható következtetéseket. Esetenként azonban elkerülhetetlen az egyes problémák felvetése. E támakörrel kapcsolatos, még nem vizsgált pszichológiai kérdések kísérleti vizsgálata különálló (célkísérlettel megoldható) feladat, amit e dolgozat csak megalapozhat.

A szubjektív valószínűségeloszlás felmérése

Alapvető feladat a q_i számok meghatározása. Kísérleteinknél általában az általunk „fogadásos módszernek" nevezett eljárást alkalmaztuk. Mindenekelőtt a lehetséges válaszokat, tehát a teljes szubjektív eseményrendszert kell meghatározni. Ez általában könnyű feladat, sőt legtöbbször eleve adott. Ezután 100 vagy más, a kísérletnek megfelelő számú „szavazási lehetőséget" adunk a kísérleti személynek valamilyen formában, és megkérjük, fogadjon az egyes eseményekre úgy, hogy a leadott szavazatok száma arányos legyen az egyes válaszok, illetve események bekövetkezésének esélyeivel. Azt is megtehetjük, hogy a leadott összes szavazatok számát nem határozzuk meg, csupán az utóbbi kikötést tesszük. Ha az egyes eseményekre tett szavazatok száma rendre $k_1, k_2 \ldots, k_n$ és ezek összege N, akkor

$$q_i = \frac{k_i}{N} \qquad (i = 1, 2, \ldots, n).$$

Különösen fontos, hogy a kisérleti személy megértse, mit jelent a szavazatok arányos elosztása. Kísérleteinknél ezt több konkrét példa bemutatásával igyekeztünk elérni.

A szubjektív valószínűségeloszlás felmérésével kapcsolatban Cohen, Dearnaley és Hansel (1956) azt a kérdést vizsgálták, a külön-külön felmért szubjektív valószínűségek összege teljes eseményrendszer esetén 1-et ad-e? Megállapították, hogy ez nem minden esetben teljesül. (Saját vizsgálatainkban a fent leírt módszer biztosítja, hogy a szóbanlevő összeg értéke mindig 1.) Ezek alapján azonban felvethető a kérdés, hogy a felmérés módszere befolyásolja-e a szubjektív valószínűség értékét. Véleményünk szerint a szubjektív valószínűséget nem egy konkrét számmal, hanem egy számhalmaz (pontosabban egy valószínűségi változó) alapján kell mérni. A felmérés módszerétől is függ, hogy e számhalmazból melyik érték kerül előtérbe, melyiket hozza a kísérletvezető tudomására a kisérleti személy E számhalmaz pontosabb leírására visszatérünk. Ez előtt néhány példát mutatunk be végrehajtható kísérletekre.

A végrehajtható kísérletek csoportosítása

A kísérleteket először két nagy csoportra osztjuk:

A) A kérdés arra vonatkozik, hogy egy véletlen eseményhalmazból melyik következik be (tehát s kísérletvezető sem tudja előre a helyes választ.

B) A kérdés bekövetkezett, tehát a kísérletvezető számára ismert eseményre vonatkozik.

A) Tekintsünk egy valószínűségszámítási értelemben vett kísérletet, amelynek $A_1, A_2, \ldots, A_n$ eseményei teljes eseményrendszert alkotnak. Ezek matematikai valószínűségei rendre $p_1, p_2, \ldots, p_n$. Ha a már leírt módon felmérjük az ezen eseményre vonatkozó szubjektív valószínűségeket, akkor a $q_1, q_2, \ldots, q_n$ értékeket kapjuk.

Az ilyen típusú kísérleteket további két csoportra oszthatjuk: az első csoportba tartoznak azok a kísérletek, amelyeknél a matematikai eloszlás meghatározható a feltételekből, a másodiknál ez nem áll, hanem statisztikai mérések vagy korábbi ismeretek szükségesek (a kísérletvezető számára is) a matematikai valószínűségeloszlás meghatározásához.

1. A matematikai eloszlás meghatározható a feltételekből.

a) Diszkrét eset: Itt azt a példát említjük, amit az Egy speciális tanulási folyamat vizsgálata információelméleti módszerekkel című dolgozatban már leírtunk (Magyar Pszichológiai Szemle 1976/5.), ahol e kísérlet más vonatkozásait vizsgáltuk.

A kísérlethez n számú doboz és N számú golyó szükséges. A golyók közül $N - 1$ fehér, egy pedig piros színű. A golyókat véletlenszerűen helyezzük el a dobozokba úgy, hogy az i-edik dobozba s_i számú ($i = 1, 2, \ldots, n$) golyó jusson $\left(\sum_{i=1}^{n} s_i = N\right)$. Annak matematikai valószínűsége, hogy a piros golyó az i-edik dobozba kerül $p_i = \frac{s_i}{N}$, ami a fenti adatok ismeretében meghatározható.

Fontos, hogy a kisérleti személy lássa a véletlenszerűséget, azt, hogy a kísérletvezető nem befolyásolhatta a piros golyó helyét.

A kisérleti személy elé azt a feladatot tűzzük, hogy fogadjon az egyes dobozokra, tehát azt kell eldöntenie, hogy milyen eséllyel kerülhet a piros golyó az egyes dobozokba.

b) Folytonos eset: Talán a folytonos valószínűségi változóra vonatkozó kísérlet megtervezése és elemzése a legnehezebb feladat, mert elég nehéz alkalmas kísérletet találni, nehezebb megértetni a kisérleti személyel a feladatát, és mélyebb matematikai ismeretekre van szükség az elemzéshez.

Példaként egy egészen egyszerű kísérletet írunk le. Ha egy kör alakú céltáblára véletlenszerűen (célzás nélkül) lövünk, és feltételezzük, hogy a céltáblát eltaláljuk, akkor valószínűségi változónak tekinthetjük azt, hogy a golyó a kör középpontjától hány egységnyire csapódik be. Ha még azt is feltételezzük, hogy a terület nagyságával arányos annak valószínűsége, hogy egy adott tartományba esik a lövés, akkor a bevezetett valószínűségi változó eloszlásfüggvénye az

$$F(x) = \begin{cases} 0, & \text{ha} \quad x < 0 \\ \dfrac{x^2}{r^2}, & \text{ha} \quad 0 \leq x \leq r \\ 1, & \text{ha} \quad x > r \end{cases}$$

függvény, ahol r a céltábla sugarát jelenti (Prékopa, 1974).

Feladat itt a szubjektív eloszlásfüggvény vagy sűrűségfüggvény meghatározása. Ennek legegyszerűbb módja az, hogy a szóbanlevő intervallumot, ahol a valószínűségi változó felveszi értékeit (jelen esetben a [0; r] intervallumot), valahány egyenlő részre osztjuk, és az egyes részekre szavazatatunk a kisérleti személlyel. A beosztás finomításával a kívánt pontossággal rajzolhatjukfel a szubjektív eloszlásfüggvényt.

Folytonos valószínűségi változó esetén gondolhatunk arra is, hogy magát a szubjektív eloszlás- vagy sűrűségfüggvényt rajzoltatjuk fel a kisérleti személlyel. Ez azonban csak magas intelligenciaszintű személyek esetén lehetséges, hisz ilyen formában nehéz megérteni a feladatot.

2. A matematikai eloszlás nem határozható meg a feltételekből.

Ha azt a kérdést vetjük fel, hogy egy adott területen elültetett diófa normális körülmények között mennyi ideig fog élni, akkor a válasz nem lehet határozott. Az élettartamot megadhatjuk matematikai valószínűségeloszlással, ez azonban a feltételekből még nem állapítható meg, ismerni kell a diófa élettartamára vonatkozó statisztikai adatokat. A szubjektív eloszlást bizonyos részintervallumokra (pl. 5 évenként) a már leírt módon határozhatjuk meg. A szubjektív és a matematikai eloszlás eltérése attól függ, hogy a kisérleti szeszemély mennyire ismeri az említett statisztikai adatokat, a diófa átlagos élettartamát. Lényegében tehát ismeret felméréséről van szó.

Ha a kérdést úgy vetjük fel, hogy mennyi ideig fog élni a diófa, akkor az idő mint folytonos valószínűségi változó szerepel, ha viszont azt kérdezzük, hány évig fog élni, akkor diszkrét esettel állunk szembe.

B) Az előbbi kísérletekre az volt a jellemző, hogy a kísérletvezető sem tudta előre, melyik esemény fog bekövetkezni. Ez a következőkre nem teljesül. Olyan kísérletekre gondolunk, ahol csak a kisérleti személy számára bizonytalan az esemény, tehát a válasz; így matematikai eloszlásról nem beszélhetünk, csak szubjektívről. (Elfajult matematikai eloszlásról szólhatnánk, de ennek nincs különösebb jelentősége számunkra.)

Diszkrét estre a következő példát említjük: Egy körlapra pl. 87 pontot rajzolunk, és megkérdezzük a kísérleti személytől, hogy szerinte hány pont van felrajzolva. A válaszra adott lehetőségek (melyekre a kísérleti személy fogad) a következők lehetnek: 0—20 között, 20—40 között, . . ., 180—200 között, 200-nál több.

Teljesen hasonló példát említhetünk folytonos esetre is. Egy bizonyos távolságban elhelyezünk egy adott hosszúságú tárgyat, és a kisérleti személynek a tárgy hosszára kell fogadnia. A szubjektív eloszlás meghatározásához az előbbi módszert itt is alkalmazhatjuk. Ezeket a kísérleteket becsléseknek is nevezhetjük.

A szubjektív valószínűségeloszlás jellemzése

Bevezettük a szubjektív valószínűségeloszlás fogalmát, leírtuk annak felmérési lehetőségét. E fogalmat azonban a fentiekkel még nem tisztáztuk eléggé. Már felhívtuk a figyelmet arra a problémára, hogy a szubjektív valószínűséget nem jellemezhetjük konkrét számmal. Kísérleteinknél beigazolódott, hogy csak a legritkább esetben van a kisérleti személynek határozott véleménye a fogadást illetően. Az egyik körgyűrűs kísérletnél például hat fogadási lehetőséget adtunk, és volt olyan személy aki a következő módon fogadott:

Körgyűrűk	1.	2.	3.	4.
1. fog.	50	25	20	5
2. fog.	5	20	25	50
3. fog.	10	20	35	35
4. fog.	5	20	33	42
5. fog.	5	10	25	60
6. fog.	2	10	44	44

Az ilyen mérvű bizonytalanság nem általános, de létezésével számolnunk kell, tehát a szubjektív valószínűség felmérése mellett annak stabilitását is vizsgálni kell.

Elgondolásunk szerint ha szubjektív valószínűségeloszlásról van szó, akkor egy eloszláshalmazra kell gondolnunk, azokra az eloszlásokra, amelyeket a kísérleti személy elvben számításba vehet. A kisérleti személy által leírt konkrét eloszlást is szubjektív eloszlásnak nevezzük, de a két fogalom nem keverhető össze.

A szubjektív valószínűségeloszlás felmérése tehát akkor lesz teljesebb, ha felmérjük, hogy a kísérleti személy milyen eloszlásokat tart lehetségesnek. Ennek egyik módja az, hogy több (legjobb talán, ha határozatlan számú) fogadási lehetőséget adunk.

Tegyük fel, hogy m számú fogadást tett a kisérleti személy q_{ij} jelentése az i-edik ($i = 1, 2, \ldots, n$) eseményre vonatkozó j-edik ($j = 1, 2, \ldots, m$) fogadásból számított szubjektív valószínűséget, Az A_i eseményhez tehát a $q_{i1}, q_{i2}, \ldots, q_{im}$ szubjektív valószínűségértékek tartoznak.

Esetenként könnyebben lehet megfogalmazni mondanivalónkat, ha egy rögzített $q_{1j}, q_{2j}, \ldots, q_{nj}$ szubjektív eloszlást az n-dimenziós tér egy pontjának tekintjük és Q_j-vel jelöljük, ahol q_{ij} számok ($i = 1, 2, \ldots, n$) a pont koordinátái. Ilyen szemléletmód szerint a szubjektív eloszlást egy n-dimenziós ponthalmazzal jellemezhetjük.

Az eloszláshalmaz áttekintése sokszor nehézkes, ezért bevezetjük az átlagos szubjektív valószínűségeloszlás fogalmát. Egy A_i esemény átlagos szubjektív valószínűségén az A_i-re felmért egyes szubjektív valószínűségek számtani átlagát értjük, és $\bar{q}_i$-vel jelöljük. Ha tehát a kísérleti személy m számú fogadást tett az eseményrendszerre, akkor

$$\bar{q}_i = \frac{\sum_{j=1}^{m} q_{ij}}{m} \quad (i = 1, 2, \ldots, n).$$

Az átlagos szubjektív valószínűségeloszlás az n-dimenziós térben az előbb leírt Q_i pontok súlypontjának felel meg.

Az elvégzett kísérleteknél általában nem a most leírt módon határoztuk meg az átlagos szubjektív valószínűségeloszlást, mert úgy láttuk, hogy reálisabb képet kapunk, ha több, kb. azonos intelligenciaszintű kísérleti személyen felmért szubjektív valószínűségek átlagát vesszük. A fenti képlet ebben az esetben is alkalmazható, csak a q_{ij} jelentését kell átfogalmazni. Ekkor q_{ij} a j-edik kísérleti személy által megadott, A_i-re vonatkozó szubjektív valószínűséget jelöli.

Az előbbinél hasznosabbnak mutatkozik az a szemléletmód, ha a $q_{i1}, q_{i2}, \ldots, q_{im}$ számokat egy q_i valószínűségi változó értékeinek tekintjük. Ez indokolt, mivel e számok nem határozott véleményt fejeznek ki, egy bizonyos fokig véletlen, hogy éppen ezeket adja meg a kisérleti személy. A q_{ij} szubjektív valószínűséghez, mint a q_i valószínűségi változó konkrét értékéhez, újabb valószínűségértéket rendelhetünk, ami azt fejezi ki, mi a valószínűsége annak, hogy éppen ezt fogja megadni a kisérleti személy. Az egyszerűség kedvéért tegyük fel, hogy e számok (szubjektív) valószínűségei egyenlőek, tehát minden q_{ij} szám valószínűsége $\frac{1}{m}$. (Lenne rá mód, hogy a szubjektív valószínűség szubjektív valószínűségét pontosabban is meghatározzuk, azonban ez már feleslegesen bonyolítaná a kísérlet kiértékelését.)

Ilyen szemléletmód mellett $\bar{q}_i$ nem más, mint a q_i valószínűségi változó várható értéke. A $\bar{q}_i$ várható érték magában nem sokat mond az A_i-hez tartozó szubjektív valószínűségről. Nem mindegy, hogy a q milyen intervallumon veszi fel az értékeit, és ezen belül hogyan helyezkednek el ezek az értékek. A q_i eloszlásának másik igen fontos jellemzője a szórás mértéke, ami lényegében a szubjektív valószínűség stabilitását is méri. Ha a q_i szórását d_i-vel, a relatív szórást r_i-vel jelöljük, akkor az alábbi összefüggéseket írhatjuk fel:

$$d_i = \sqrt{\frac{1}{m} \cdot \sum_{j=1}^{m} (q_{ij} - \bar{q}_i)^2}, \quad r_i = \frac{d_i}{\bar{q}_i}.$$

Azoknál a kísérleteknél, amelyeknél létezik matematikai valószínűségeloszlás, szükséges lehet annak megvizsgálása, hogy a szubjektív eloszlás mennyiben tér el a matematikaitól. Legegyszerűbb, ha a matematikai és az átlagos szubjektív eloszlás távolságával mérjük az eltérést. Ez lényegében két n-dimenziós pont távolságának felel meg, tehát a

$$T = \sqrt{\sum_{i=1}^{n} (p_i - \bar{q}_i)^2} \text{ képlettel számítható ki.}$$

Folytonos esetben az empirikus szubjektív sűrűségfüggvény és a matematikai sűrűségfüggvény eltérésével is számolhatunk.

A szubjektív valószínűségeloszlás annál jobb, minél közelebb áll a matematikaihoz. Mivel e két eloszlás maximális távolsága $\sqrt{2}$, ezért a szubjektív eloszlás „jóságát" (J) célszerű az alábbi képlettel számítani:

$$J = \frac{\sqrt{2} - T}{\sqrt{2}}:$$

J a [0;1] intervallumon veszi fel értékeit. Ha a szubjektív és a matematikai eloszlás megegyezik, akkor $J = 1$, ha a kettő távolsága maximális, akkor $J = 0$.

Itt csak a kísérlet kiértékelésével közvetlenül kapcsolatban levő mennyiségekkel foglalkoztunk, melyek a szubjektív eloszlás meghatározására és a szubjektív valószínűség stabilitásának mérésére irányulnak. Az eloszlást jellemző további fogalmak (a szubjektív várható érték, a szubjektív eloszlás szórása stb.) a valószínűségszámításban szokásos módon vezethetők be.

A végrehajtott kísérletek elemzése

Az általunk végzett kísérleteknek kettős célja volt. Egyrészt azt vizsgáltuk, hogy az elméleti meggondolások reális alapokon nyugszanak-e, és a leírt matematikai apparátus alkalmas-e a szubjektív valószínűségeloszlás vizsgálatára.

Másik célunk a problémafelvetés, illetve a figyelem ráirányítása az egyes kérdésekre. E célok érdekében több, különböző témájú kísérletet végeztünk.

E kísérletek lényegében három téma köré csoportosulnak, melyeket a problémák jobb megvilágítása érdekében több módon és több csoportban is elvégeztünk. Az alábbiakban többnyire a szakközépiskola harmadik osztályának 18 tanulójával végzett kísérletek adatait közöljük, de esetenként (ezeket külön jelezzük) utalunk a középiskola esti tagozatára járó 12 végzős hallgatóval végzett kísérletek eredményeire is. A tanulókkal végzett kísérletek kiértékelésekor három tizedesjegy pontosságig (a harmadik jegy kerekített érték), a felnőtteknél csak két tizedesjegy pontosságig számoltunk.

Első témakör a már leírt „dobozos" kísérlet. Ezeknél a kísérleteknél a matematikai valószínűségeloszlás létezik és a feltételekből könnyen meghatározható. Valamennyi ide tartozó kísérletet úgy végeztük el, hogy minden kísérleti személynek csak egy fogadási lehetőséget adtunk, tehát $\overline{q}_i$ a több kísérleti személy által leírt szubjektív valószínűségek átlagát (várható értékét) jelenti. Többnyire a szubjektív és a matematikai eloszlás távolságát és a szubjektív eloszlás jóságát figyeltük, de esetenként kiszámítottuk a szubjektív valószínűség szórását is, amiből annak stabilitására következtethetünk.

1. kísérlet: Három dobozba véletlenszerűen osztottunk el 40 golyót (39 fehér és 1 piros) úgy, hogy az egyes dobozokba 5, 25 és 10 golyó került. A golyók elosztása a kisérleti személyek előtt történt, de a golyók színét nem láthatták. A jobb áttekinthetőség érdekében az egyes dobozokra ráírtuk a bennük levő golyók számát. A kisérleti személyeknek a piros színű golyó helyére kellett fogadniok, a rendelkezésükre álló 100 szavazási lehetőséget kellett arányosan elosztani.

A kísérlet eredményeit az alábbi táblázat tartalmazza:

i	1	2	3
p_i	0,125	0,625	0,25
$\overline{q}_i$	0,163	0,537	0,3
d_i	0,044	0,065	0,037
r_i	0,271	0,122	0,124

$T = 0{,}108 \quad J = 0{,}924$

p_i annak matematikai valószínűségét jelöli, hogy a piros golyó az i-edik dobozba kerül.

Ha e kísérlet eredményeit összehasonlítjuk a következő három kísérlet eredményével, akkor azt tapasztaljuk, hogy itt a legnagyobb a szubjektív és a matematikai eloszlás távolsága annak ellenére, hogy ez a legegyszerűbb kísérlet. Ezt annak tulajdoníthatjuk, hogy a végrehajtás sorrendjében is ez volt az első kísérlet, és a próbakísérlet ellenére is gyakorlatlanok voltak a tanulók.

A szórásértékek viszont meglepően kicsik, tehát nagy a szubjektív valószínűség stabilitása.

Megjegyezzük még, hogy ennél a kísérletnél szépen kirajzolódott az az ismert pszichológiai törvényszerűség, hogy az ember a kisebb valószínűségeket növelni, a nagyobbakat pedig csökkenteni igyekszik (Rupp 1974); a matematikai eloszlás mintegy ,,kisimulva'' jelenik meg a tudatunkban.

2. kísérlet: Az előbbi kísérlettől csak annyiban különbözik, hogy összesen 13 golyót használtunk. Az első dobozba 10, a másodikba 1, a harmadikba pedig 2 golyót helyeztünk. Célunk az volt, hogy az egyenletes eloszlástól az előbbinél jobban eltérő esetben is megvizsgáljuk a szubjektív eloszlás jellemzőit.

A kapott eredmények:

i	1	2	3
p_i	0,769	0,077	0,154
$\bar{q}_i$	0,761	0,065	0,174
d_i	0,158	0,054	0,103
r_i	0,207	0,835	0,588

$T = 0{,}023 \qquad J = 0{,}983$

Ha e kísérlet eredményeit az előbbivel összehasonlítjuk, akkor két lényeges különbséget vehetünk észre. Az egyik a szubjektív és matematikai eloszlás távolságának csökkenése, amit már indokoltunk. A másik a szubjektív valószínűségek szórásainak növekedése, ami valószínűleg az eloszlás megváltozásának következménye. A q_1 valószínűségi változó szórásának növekedését azzal magyarázhatjuk, hogy a p_1 értéke nagy, és így viszonylag kisebb, de abszolút értékben nagy eltérések nagyobb szórásértékhez vezetnek. Ez abból is látható, hogy a relatív szórás értéke nem változott lényegesen az előbbiekhez viszonyítva. q_2 esetében viszont fordított a helyzet: a szórása nem sokban különbözik az előbbiek átlagától, relatív szórása viszont igen nagy. Úgy tűnik, a túl kicsi és a túl nagy matematikai valószínűségű események szubjektív valószínűségének stabilitása rosszabb, mint a közbülső esetekben.

3. kísérlet: Formailag ez is megegyezik az előbbi kísérletekkel, de itt 8 dobozt és 38 golyót használtunk.

A golyók elosztása:

1.	2.	3.	4.	5.	6.	7.	8.	doboz
6	4	10	2	4	7	3	2	golyó.

Az események nagyobb száma miatt itt nem ragaszkodtunk ahhoz, hogy pontosan 100 szavazatot adjon a kisérleti személy. Így megakadályoztuk azt, hogy az utolsó dobozokra az esetleges ,,maradék'' szavazati lehetőségeket osszák el.

Ezzel a kísérlettel az volt a célunk, hogy az események nagyobb száma mellett is megvizsgáljuk a szubjektív eloszlás jóságát.

Kapott eredmények:

i	1	2	3	4	5	6	7	8
p_i	0,158	0,105	0,263	0,053	0,105	0,184	0,078	0,053
$\bar{q}_i$	0,142	0,096	0,317	0,041	0,097	0,178	0,074	0,051

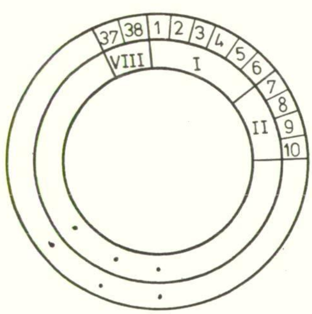
$T = 0{,}059 \qquad J = 0{,}958$

Az események számának növelésével nem változott lényegesen a szubjektív és a matematikai eloszlás távolsága, tehát a szubjektív eloszlás jósága sem.

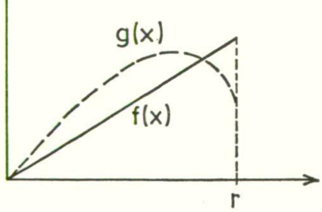
4. kísérlet: Kísérleteinknél fontosnak tartottuk annak megvizsgálását, hogy az eseményrendszer megjelenési formája hogyan befolyásolja a szubjektív eloszlást, illetve annak jellemzőit. Ezért végeztük el a következő kísérletet, amelynél a matematikai eloszlás megegyezik az előbbivel, formailag viszont lényegesen különbözik attól.

Két koncentrikus körgyűrű közül a külsőt 38 (arab számmal jelölt) egyenlő részre, a belsőt pedig 8 (római számmal jelölt) részre osztottuk úgy, hogy a belső körgyűrű egyes beosztásai az alábbi táblázat szerint „tartalmazzák" a külső körgyűrű beosztásait:

I.	II.	III.	IV.	V.	VI.	VII.	VIII.	
6	4	10	2	4	7	3	2	(1. ábra)



Egy dobozba 1-től 38-ig számozott cédulákat helyeztünk, amiből véletlenszerűen (húzással) választottunk ki egyet, de nem mutattuk meg a kísérleti személyeknek, akiknek arra kellett fogadniok, hogy a kihúzott (arab) szám melyik belső körgyűrűcikkhez tartozik.



A matematikai és a szubjektív valószínűségértékeket az alábbi táblázat tartalmazza:

i	I.	II.	III.	IV.	V.	VI.	VII.	VIII.
p_i	0,158	0,105	0,262	0,053	0,105	0,184	0,078	0,053
$\bar{q}_i$	0,137	0,085	0,3	0,055	0,099	0,195	0,073	0,056

T = 0,049 J = 0,965

A kísérlet eredményeiből megállapíthatjuk, hogy ebben az esetben a formai változás nem okozott lényeges változást a szubjektív valószínűségeloszlásban.

A leírt négy kísérlet eredményeit összevetve megállapíthatjuk, hogy ennél a kísérletcsoportnál a szubjektív eloszlás jól közelíti a matematikait. A szubjektív eloszlás jóságát nem befolyásolja lényegesen sem az események száma, sem az eseménytér megjelenési formája. Nyilvánvaló, hogy mindez a kísérletek egyszerűségéből is adódik, hisz itt könnyű átlátni az esélyeket.

Valóban valamivel rosszabb becslésekkel találkozunk a második kísérletcsoportnál, a „céltáblás" kísérleteknél, ahol ugyanannak a folytonos valószínűségi változónak az eloszlását különböző feltételek mellett becsülték meg a kisérleti személyek; majd ezeket a becsléseket összehasonlítottuk.

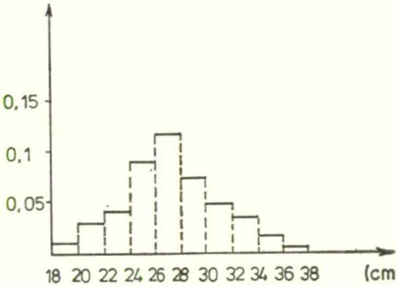
5. kísérlet: Egy körlapra négy egyenlő vastagságú körgyűrűt rajzoltunk (a belső kör sugara megegyezik a körgyűrűk vastagságával), ezeket belülről kifelé haladva sorszámoztuk. A kísérleti személynek arra kellett fogadnia, hogy 100 véletlenszerű lövésből (feltéve, hogy nem célzunk a kör középpontjába, de mindig eltaláljuk a céltáblát) hány esik az egyes körgyűrűkbe. A matematikai eloszlást itt úgy határozzuk meg, hogy az 1-et a körgyűrűk területének arányában osztjuk fel. E kísérleteknél mód adódik annak vizsgálatára, hogy a szubjektív eloszlás létrejöttében a területek aránya vagy a körgyűrűk elhelyezkedése vezérli-e a kísérleti személyt.

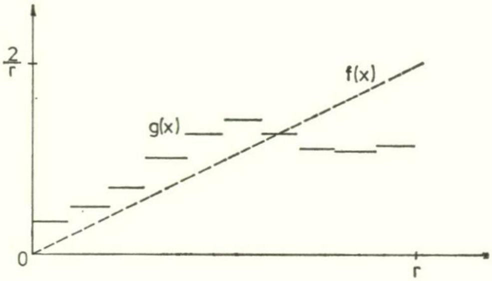
A kísérlet statisztikai adatai:

i	1	2	3	4
p_i	0,062	0,187	0,313	0,438
$\bar{q}_i$	0,138	0,242	0,342	0,279
d_i	0,05	0,071	0,103	0,097
r_i	0,361	0,294	0,301	0,348

T = 0,187 J = 0,868

Ennél a kísérletnél igen érdekes jelenség nyilvánult meg. A kör középpontjától kifelé haladva a szubjektív valószínűségek egy darabig (matematikaihoz hasonlóan) növekednek, azután hirtelen csökkennek, holott a matematikai valószínűségek tovább növekednek. Ugyanez a tendencia nyilvánult meg akkor is, ha ezt a kísérletet más csoportban hajtottuk végre, és akkor is, ha 4 körgyűrű helyett 10-et vettünk fel a céltáblára. Az elmondottak szemléltetése érdekében az utóbb említett (10 körgyűrűs) kísérlet eredményeit grafikonon mutatjuk be (2. ábra).





A folytonos vonallal rajzolt g(x) függvény az empirikus szubjektív sűrűségfüggvényt, a szaggatottal rajzolt f(x) függvény pedig a matematikai sűrűségfüggvényt ábrázolja.

E kísérletnél is felléphetett az a már említett tulajdonság, hogy az ember a kisebb valószínűséget növelni, a nagyobbat csökkenteni igyekszik, de nem ez látszik meghatározónak. Úgy tűnik, hogy egy bizonyos „széli efektus" lépett fel. Mivel feltétel volt, hogy minden lövés eltalálja a céltáblát, a kísérleti személy biztosítani akarta magát, ezért csökkentette a szélső gyűrűk esélyeit. Ennek igazolására végeztük el a következő két kísérletet:

6. kísérlet: Találomra lövünk az 5. kísérletnél leírt céltábla felé meghatározatlan lövésszámmal (a céltáblát akár le is takarhatjuk). Tegyük fel, hogy a céltáblát éppen 100 lövés találta el. A kísérleti személynek most is arra kell fogadnia, hogy a 100 lövés hogyan oszlik meg a körgyűrűk között. Látszólag a feladat megegyezik az előbbivel (a matematikai eloszlás nem is változott), megfogalmazása azonban különbözik attól, és ez lényeges hatással volt a szubjektív eloszlásra.

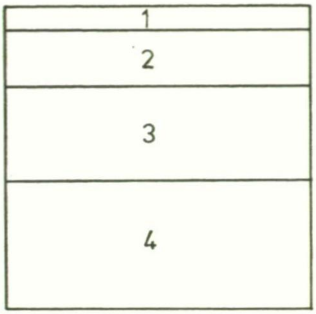
A kapott eredmények:

i	1	2	3	4
p_i	0,062	0,187	0,313	0,438
$\bar{q}_i$	0,105	0,215	0,326	0,354

$T = 0{,}099$ $\quad J = 0{,}93$

Itt már csökkent az előbbi effektus, és a szubjektív eloszlás jósága lényegesen javult.

7. kísérlet: Egy négyzet alakú céltáblát készítettünk úgy, hogy az egyik oldallal párhuzamos szakaszok az 5. kísérletnél leírt körgyűrűk területének arányában osztják fel a négyzetet (3. ábra).



A kísérleti személyeknek a 6. kísérletnek megfelelően fogalmaztuk meg a feladatot.

A kapott eredmények:

i	1	2	3	4
p_i	0,062	0,187	0,313	0,438
q_i	0,057	0,164	0,296	0,482

$T = 0,053 \qquad J = 0,963$

Mint várható volt, ennél a kísérletnél már nyoma sincs az előbbieknél tapasztalt széli effektusnak. Sőt, a szubjektív valószínűségeloszlás meglepően jól közelíti meg a matematikait. Ezt a jó közelítést annak tulajdoníthatjuk, hogy itt igen jól érzékelhető a területek aránya, ami a körgyűrűk esetében nem mondhatók el. Az előbbi kísérleteknél tehát a körgyűrűk helyzete erősen befolyásolta a szubjektív valószínűségértékeket.

A hatodik kísérletet úgy is elvégeztük, hogy a kisérleti személynek hat fogadási lehetőséget adtunk, így mód nyílt az egyéni stabilitás felmérésére is. $\bar{q}_i$ tehát az egy tanuló által adott szavazatok alapján számított szubjektív valószínűségek várható értékét, d_i pedig ezek szórását jelenti. A legkisebb és legnagyobb szórású szubjektív eloszlást olyan kisérleti személyek adatai közül választottuk ki, akik éltek mind a hat fogadási lehetőséggel.

		1	2	3	4
Legkisebb	$\bar{q}_i$	0,088	0,258	0,342	0,312
szórású	d_i	0,048	0,053	0,044	0,077
Legnagyobb	$\bar{q}_i$	0,093	0,117	0,373	0,417
szórású	d_i	0,116	0,103	0,156	0,215

Azt tapasztaltuk, hogy az egyéni stabilitás általában rosszabb, mint a több kísérleti személy alapján számított stabilitás. Nem lehet azonban véletlen, hogy általában az elsőnek leadott szavazatok alapján számított szubjektív eloszlás közelíti legjobban a matematikait, a többi ettől sokszor fokozatosan távolodik. Ez az észrevétel egyúttal igazolja azt az általánosan alkalmazott eljárásunkat is, hogy több kísérleti személytől kértünk egy-egy fogadást, és ezek átlagát vettük vizsgálatunk alapjául. Erre utaltunk az előbbi fejezetben, amikor azt mondtuk, hogy reálisabb eredményhez jutunk, ha több kísérleti személyre vonatkozó szubjektív valószínűségek átlagával dolgozunk.

Érdekes kísérletnek mutatkozott a különböző hosszúságok megbecsültetése. Mivel itt matematikailag nem véletlen eseményre kérdezünk, ezért nem beszélhetünk a szubjektív és a matematikai eloszlás távolságáról. Figyelmünket a valódi hossz és az átlagos szubjektív eloszlásból számított várható érték összehasonlítására fordítjuk. A következő adatok a felnőttek csoportjában elvégzett kísérletek kiértékeléséből származnak.

Először egy 27,7 cm hosszú egyenes vonalat rajzoltunk fel (a valódi hosszúságot természetesen nem ismerték a kísérleti személyek), majd a 100 cm-es intervallumot (először 10 cm-es, majd 2 cm-es) részintervallumokra osztottuk. A kísérleti személynek arra kellett fogadnia, hogy a felrajzolt vonal hossza melyik intervallumba esik. A fogadások alapján számított átlagos szubjektív valószínűségeloszlásokat az alábbi táblázatból olvashatjuk ki. Azokhoz az intervallumokhoz, amelyek a táblázatban nem szerepelnek, 0 szubjektív valószínűségérték tartozik.

A részintervallumok hossza 10 cm:

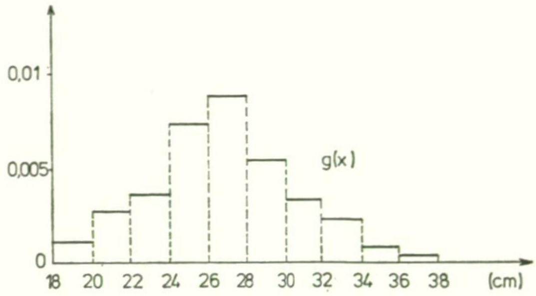
Intervallum	0—10	10—20	20—30	30—40	40—50
$\bar{q}_i$	0	0,12	0,75	0,13	0

Ha az intervallumok közepét tekintjük valószínűségi változónak, akkor a vonal hosszának szubjektív várható értéke 25,1 cm.

A részintervallumok hossza 2 cm:

Intervallum	16—18	18—20	20—22	22—24	24—26	26—28
$\bar{q}_i$	0	0,03	0,07	0,09	0,18	0,24
Intervallum	28—30	30—32	32—34	34—36	36—38	38—40
$\bar{q}_i$	0,15	0,1	0,08	0,04	0,02	0

Az előbbi feltétel mellett a szubjektív várható érték 27,3 cm.



A 4. ábrán e táblázat adatai alapján felrajzoltuk az empirikus szubjektív sűrűségfüggvényt.

Az adatokból kitűnik, hogy a két esetben felmért szubjektív várható érték majdnem megegyezik. Az ilyen kicsi különbség talán véletlen, de ha azt is figyelembe vesszük, hogy ezektől a valódi érték sem sokban különbözik, akkor határozott tendenciát állapíthatunk meg.

Érthetően a görbe vonal becslésekor a szubjektív várható érték távolabb esik a valódi értéktől és nagyobb az eloszlás szórása is. Ennél a kísérletnél 94,2 cm hosszú, erősen görbülő vonalat rajzoltunk fel, és 5 cm hosszúságú intervallumokat jelöltünk ki a fogadására. Ha az egyes kísérleti személyek által leírt szubjektív eloszlásokat összehasonlítjuk, akkor nagy eltéréseket tapasztalhatunk, tehát a szubjektív valószínűségek stabilitása kicsi. A vonal görbesége nagyon rontotta a becslés jóságát. Az átlagos szubjektív eloszlásból számított várható érték 60,3 cm-adódott. A legjobb becslést (84 cm-es várható értékkel) egy lakatos szakmunkás adta, szubjektív eloszlás szórása is kicsi a többihez viszonyítva. Ebből is látható, hogy a szubjektív eloszlás jellemzői kapcsolatba hozhatók más egyéni tulajdonságokkal is.

Összefoglalva elmondhatjuk, hogy már az általunk elvégzett kísérletek is felvetettek néhány pszichológiai jellegű problémát. Azon túlmenően, hogy a felvetett és a még kínálkozó problémák vizsgálata igen érdekes feladatnak látszik, ez a szemléletmód hasznos lehet más pszichológiai feladat megoldásánál is.

A közlemény a szerkesztőségbe érkezett: 1976. XII. 16.

IRODALOM

1. John Cohen, 1960. Chance, Skill, and Luck. U.S.A.: Penguin Books Inc.
2. John Cohen, E. J. Dearnaley and C.E.M. Hansel, 1956. The Addition of Subjective Probabilities. Acta Psychol. 12.
3. N. T. Feather, 1959. Subjective Probability and Decision Under Uncertainty. Psychological Review 66.
4. Prékopa András, 1974. Valószínűségelmélet. Műszaki Könyvkiadó, Budapest.
5. Dr. Rupp Mária, 1974. Kísérlet a véletlen átélésének pszichológiai vizsgálatára. Akadémiai Kiadó, Budapest.
6. Vaszkó Mihály—Kaucsek György, Szignál és valószínűségi bizonytalanság hatása a szimulációs szituációban végzett operátori tevékenységre. Ergonómia 1970/4.

РАСПРЕДЕЛЕНИЕ СУБЬЕКТИВНЫХ ВЕРОЯТНОСТЕЙ ПРИ РЕШЕНИИ ЗАДАЧ РАЗЛИЧНОГО ТИПА

ИШТВАН СЕКЕРЕШ

Нередко встречается, что на какой-либо вопрос невозможно дать однозначный ответ, напротив, предполагается возможными дать несколько ответов. Предполагаемые возможными ответы A_3, A_4, . . ., A_n назовем полной субъективной системой событий. При этом рапределение субъективных вероятностей можно определить таким образом, чтобы удовлетворить математическим требованиям, предельявляемым к распределению вероятностей. Определенные подобным образом вероятности находятся в пропорциональной зависимости с силой стимулов, имплицирующих отдельные ответы. Основными задачами при этом являются:. измерение субъективных вероятностей, их распределения, а также оценка результатов измерения, что демонстрируется в данной работе на основе целого ряда конкретных примеров и проведенных опытов. Основной целью опытов было стремление подтвердить существование распределения субъективных вероятностей, а также исследование стабильности субъективных вероятностей. В процессе проведения опытов и оценки полученных результатов возникли новые психологические проблемы. Наиболее интересным является то, в какой степени формальная сторона системы событий влияет на распределение субъективных вероятностей.

SUBJECTIVE PROBABILITY DISTRIBUTION IN DIFFERENT TYPES OF TASK SITUATIONS

SZEKERES, ISTVÁN

It happens that no unequivocable answer can be given to a question but several alternatives can be possible. Answers A_1, A_2, ..., A_n that we find possible are called as full subjective set of events. Fitting this a subjejective probability distribution can be defined in a way that will meet the mathematical demands laid against probability distribution. Probabilities defined this way are in proportion to the strenght of the stimuli urging to produce the certain answers.

The basic task that is the measurement of the subjective probability distribution and the evaluation of the results of measurements is demonstrated by several concrete examples and the performed experiments. The main goal of the experiments carried out was to confirm the existence of subjective probability distribution and to study its stability. Further psychological problems arise in the course of performing the experiments and evaluating the results. The question that to what extent does the formal side of the set of events influence the subjective distribution, was the most interesting one of the problems.

A SKIZOFRÉNIA A DINAMIKUS RAJZVIZSGÁLAT TÜKRÉBEN

HÁRDI ISTVÁN
Pest Megyei Tanács Ideggondozó Intézete

500 eset kapcsán skizofréniások dinamikus rajzvizsgálatáról számol be a szerző. Sorozatosan összehasonlító követés céljából a betegekkel embert rajzoltat. A kórképet az *eljárás mozgásában, elsősorban időbeliileg szemléli.* Ennek alapján „korai" és „késői" krónikus képeket lehet elkülöníteni. Amíg egyfelől a kronicitás a rajzok klisészerű ismétlődésében, szinte fotokópiaszerű másolásban mutatkozik, addig a korai képek — a reverzibilitás fokától függően — az átmeneti kóros jelenségek mutatkozása után ismét az eredeti szintre térnek vissza. A sorozatos összehasonlító követés mellett jelentős a sztereotípnek mondható ismétlődő rajzmód mellett az acut, nagy intenzitású állapotokat jellemző, váratlan, hirtelen változás, mely színes sokarcúságával konfrontációnak nevezett eljárással különíthető el: az akut állapotban készült rajzokat szembeállítja a többi taggal. A dinamikus rajzvizsgálat a személyiségen keresztül fogja fel a kórképet: mennél inkább átalakult, „skizofrenizálódott" a személyiség, annál inkább észlelhető ez a rajzban. Az elmondottak mellett még külön érdekessége van az üres és kitöltött alakoknak, melyek az autizmust és az abból való kikerülést, a világhoz való readaptációt és reszocializálódást jelentik. A dinamikus rajzvizsgálat az egyes skizofréniás alcsoportokba is betekintést nyújt.

A skizofrénia képi kifejeződése kezdettől fogva a legjobban foglalkoztatta a terület kutatóit. Így pl. az úttörő Prinzhorn munkájában (30) 75%-ban szerepeltek skizofréniás alkotások. Morgenthaler munkássága (22, 35) ugyancsak rendkívül jelentős, elsősorban a híres Wölfli esetének feldolgozásával. Bürger-Prinz (5) Bumke kézikönyvében írt fejezete e betegek szimbolikára, sztereotípiára, ornamentikára és primitívségre való készségét emeli ki. Hangsúlyozza a vizsgálódásokban a hosszmetszeti és keresztmetszeti szemlélet szükségességét. A rajz vagy festmény jellege függ a betegség időbeli szakaszától. Kretschmer (18) a stilizáló tendencia mellett képagglutinációról beszél, Bürger-Prinz ebben a vonatkozásban kontaminációról, Rennert (31, 32) pedig sűrítésről. Utóbbi, munkásságának újabb szakaszában, kiemelten elemzi a sztereoszkópos látás zavarát a skizofrének mélység nélküli, lapos, hibás perspektívákkal dolgozó ábrázolásában (33). A gyermekrajzokban ismert, hasonló jelenségekre utalva mindezt regresszívnek véli. Suchenwirth (34) az észlelt grafikus jelenségeket nem tartja specifikusnak, megfigyelései szerint a leírtak más kórképnél is előfordulhatnak. Ő is — Rennerthez hasonlóan — foglalkozik a kezelés nyomán jelentkező változásokkal. Ez Marinow (20, 21), Gamna (6), Navratil (23, 24, 26) és Brade és munkatársai munkájában (4) még kifejezettebb. Navratil anyaga főként emberrajzokat tartalmaz. Skizofréniáról szóló értékes monográfiája (23) és más közleményei (24, 26) több ponton találkoznak megfigyeléseimmel. E betegeknél Jakab-hoz (17) hasonlóan a határvonal túl-